407
8.9.70

407

THE AMERICAN
JOURNAL OF PSYCHOLOGY

EDITED BY
KARL M. DALLENBACH
UNIVERSITY OF TEXAS

AND

M. E. BITTERMAN
BRYN MAWR COLLEGE

E. B. NEWMAN
HARVARD UNIVERSITY

WITH THE COÖPERATION OF

E. G. BORING, Harvard University; W. K. ESTES, Stanford University; J. P. GUILFORD, University of Southern California; HARRY HELSON, Kansas State University; E. R. HILGARD, Stanford University; FRANCIS J. IRWIN, University of Pennsylvania; F. NOWELL JONES, University of California, Los Angeles; G. L. KREEZER, Washington University; D. G. MARQUIS, Massachusetts Institute of Technology; ARTHUR W. MELTON, University of Michigan; W. C. H. PRENTICE, Wheaton College; T. A. RYAN, Cornell University.

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY

Vol. LXXX
MEZES HALL, THE UNIVERSITY OF TEXAS
AUSTIN, TEXAS 78712

Copyright 1967 by Karl M. Dallenbach

Bureau Edni. [illegible] Research
DAVID HA [illegible] ING COLLEGE
Dated 8. 9. 70
Acc. No. 9 - 407

TABLE OF CONTENTS

ARTICLES AND NOTES

Alperson, B., and Bakan, P., Pronounceability, Attensity, and Interference in the Color-Word Test ... 416
Anderson, J. P., Fourteenth Annual Meeting of the Southwestern Psychological Association ... 645
Anstis, S. M., and Atkinson, J. Distortions in Moving Figures Viewed through a Stationary Slit ... 572
Anstis, S. M., and Loizos, C. M., Cross-Modal Judgments of Small Holes .. 51
Atkinson, J., and Anstis, S. M., Distortions in Moving Figures Viewed through a Stationary Slit ... 572
Bakan, P., and Alperson, B., Pronounceability, Attensity, and Interference in the Color-Word Test ... 416
Beck, J., and Shaw, W. A., Ratio-Estimation of Loudness-Intervals ... 59
Bevan, W., and Dukes, W. F., Stimulus-Variation and Recall: The Role of Belongingness ... 309
Birnbaum, I. M., Discrimination-Reversal, Extinction, and Acquisition after Different Amounts of Training ... 363
Bitterman, M. E., Gonzalez, R. C., and Holmes, N. K., Extinction in the Goldfish as a Function of Frequency and Amount of Reward ... 269
Blanton, R. L., and Odom, P. B., Rule Learning in Deaf and Hearing Subjects ... 591
Blessing, W. W., Coltheart, M., and Landauer, A. A., The Effect of False Perspective Cues on Distance- and Size-Judgments: An Examination of the Invariance Hypothesis ... 250
Boitano, J. J., and Isaacson, R. L., Effect of Variation in Shock-Intensity on the Behavior of Dorsal-Hippocampectomized Rats in Two Passive-Avoidance Situations ... 73
Brent, S. B., Linguistic and Non-Linguistic Processes in Serial Learning and Retroactive Interference ... 133
Britten, R., The Size-Distance Invariance Hypothesis ... 640
Browning, D., The Fifty-Ninth Annual Meeting of the Southern Society for Philosophy and Psychology ... 466
Burnham, C. A., and Worchel, S., Reduction of Autokinesis with Information about the Registration of Eye-Position ... 434
Campbell, D., Chapman, N., and Griffiths, E., An Apparatus for Detecting and Monitoring Movement ... 438
Carmichael, L., The 1966 Meeting of the American Philosophical Society 464
Carmichael, L., The 1967 Meeting of the National Academy of Sciences .. 465
Caul, W. F., and Ludvigson, H. W., Spatial Generalization as a Function of Variation in a Nonspatial Attribute, Instructions, and Individual Strategies ... 186
Ceraso, J., The Recall of Long and Short Stimulus-Lists ... 221
Chapman, N., Campbell, D., and Griffiths, E., An Apparatus for Detecting and Monitoring Movement ... 438
Cline, M. G., The Perception of Where a Person Is Looking ... 41
Coffman, T. L., Karlins, M., Lamm, H., and Schroder, H. M., Conceptual Complexity, Recall, and Identification ... 631
Coltheart, M., Blessing, W. W., and Landauer, A. A., The Effect of False Perspective Cues on Distance-and Size-Judgments: An Examination of the Invariance Hypothesis ... 250
Cooper, L. M., Hilgard, E. R., Morgan, A. H., and Voevodsky, J., The Measurement of Suprathreshold Pain ... 124

Crovitz, H. F., The Form of Logical Solutions 461
Crowder, W. F., and Pickens, R. W., A Recorder of Locomotor Activity .. 442
Dallenbach, K. M., An Acknowledgment 318
Dallenbach, K. M., An Editorial Announcement 491
Dallenbach, K. M., Errata 318, 469
David, K. H., and Dielman, T. E., Temporal Arrangement and Specificity of Set in an Autokinetic Situation 426
Dawson, W. W., and Jones, Q. R., Jr., A Dynamic Cue of the Magnitude of Tactile Sensation 287
Day, R. H., and Wade, N. J., Tilt and Centrifugation in Changing the Direction of Body-Force 637
Des Lauriers, G., Mitchell, D., and Wong, R., Serial Learning of Children as Facilitated by Mediation 140
Diamond, L., The Effect of Training Procedures on the Double Alternation-Behavior of Laboratory Rats in a Temporal Maze 594
Dickson, D., Leibowitz, H. W., and Pollard, S. W., Monocular and Binocular Size-Matching as a Function of Distance at Various Age-Levels 263
Dielman, T. E., and David, K. H., Temporal Arrangement and Specificity of Set in an Autokinetic Situation 426
Dockrill, F. J., Levy, B. A., and Mc Nulty, J. A., The Subthreshold Perception of Stimulus-Meaning 28
Donnelly, M. B., and Rimoldi, H. J. A., Experimental Considerations Concerning Category and Magnitude Scaling 602
Dukes, W. F., and Bevan, W., Stimulus-Variation and Recall: The Role of Belongingness 309
Eagle, M. N., The Effect of Learning Strategies upon Free Recall 421
Epstein, W., Perceptual Learning Resulting from Exposure to a Stimulus-Invariant 205
Eyman, R. K., The Effects of Sophistication on Ratio- and Discriminative Scales 520
Fisher, G. H., A Common Principle Relating to the Müller-Lyer and Ponzo Illusions 626
Fisher, G. H., Distortion Gradients and the Necker Cube 453
Fisher, G. H., Measuring Ambiguity 541
Fitzgerald, R. E., and Marshall, A. J., Left-Right Field Differences with Partial Report of Letters 370
Flax, S. W., and Hahn, W. W., A Solid-State Shock Absorber 446
Freeman, R. B., Jr., Optical Calibration of the Frontal Plane 302
Galloway, A., and Kennedy, A., Retrieval Time and the Effect of Serial Position 430
Gilbert, A. J., Factors Involved in Judgments Made by the Method of Constant Stimuli 586
Glick, J. A., Rand, G., Wapner, S., and Weinberg, J., Effect of Speed of Movement on Tactual Kinesthetic Perception of Extent 608
Glickman, M., and Segal, S. J., Relaxation and the Perky Effect: The Influence of Body Position on Judgments of Imagery 257
Gonzalez, R. C., Bitterman, M. E., and Holmes, N. K., Extinction in the Goldfish as a Function of Frequency and Amount of Reward 269
Griffiths, E., Campbell, D., and Chapman, N., An Apparatus for Detecting and Monitoring Movement 438
Grissom, R. J., Group Experiment in Hearing 120
Gumenik, W. E., and Webb, D., Effects of the Relevance of Interpolated Stimuli on Time-Error in the Comparative Judgment of Weights 115
Gunn, W. J., and Loeb, M., Correlations of Performance in Detecting Visual and Auditory Signals 236
Hahn, W. W., and Flax, S. W., A Solid-State Shock Absorber 446
Hall, J. F., Relationships among Number of Measures of Meaningfulness .. 291
Hayes, D. P., and Meltzer, L., Bone-Conducting Microphones 619

Helson, H., and Woodruff, B., Torque-Sensitivity as a Function of Knob-Radius and Load 558
Hepburn, K. L., Pappas, B. A., and Suboski, M. D., A Comparison of Recall and Binary Recognition 398
Hilgard, E. R., Kenneth Wartenbee Spence: 1907-1967 314
Hilgard, E. R., Cooper, L. M., Morgan, A. H., and Voevodsky, J., The Measurement of Suprathreshold Pain 124
Hirsh, I. J., Max Frederick Meyer: 1873-1967 644
Hofman, J. E., An Analysis of Concept-Clusters in Semantic Inter-Concept Space 345
Holmes, D. S., 'Closure' in a Gapped Circle Figure 614
Holmes, N. K., Bitterman, M. E., and Gonzalez, R. C., Extinction in the Goldfish as a Function of Frequency and Amount of Reward 269
Irwin, R. J., and Porsolt, R. D., Binaural Summation in Loudness of Two Tones as a Function of Their Bandwidth 384
Isaacson, R. L., and Boitano, J. J., Effects of Variation in Shock-Intensity on the Behavior of Dorsal-Hippocampectomized Rats in Two Passive-Avoidance Situations 73
Iverson, M. A., The Thirty-Eighth Annual Meeting of the Eastern Psychological Association 468
Jarrard, L. E., and Lewis, T. C., Effects of Hippocampal Ablation and Intertrial Interval on Acquisition and Extinction of a Complex Maze 66
Jerison, H. J., Apparent Motion of a Vista: An Illusion of Perspective 448
Jones, Q. R., Jr., and Dawson, W. W., A Dynamic Cue of the Magnitude of Tactile Sensation 287
Judisch, J. M., and Leibowitz, H. W., The Relation between Age and the Magnitude of the Ponzo Illusion 105
Judisch, J. M., and Leibowitz, H. W., Size-Constancy in Older Persons: A Function of Distance 294
Julius, M., Maier, N. R. F., and Thurber, J. A., Studies in Creativity: Individual Differences in the Storing and Utilization of Information .. 492
Karlins, M. Coffman, T. L., Lamm, H., and Schroder, H. M., Conceptual Complexity, Recall, and Identification 631
Katz, M. S., Feedback and Accuracy of Target Positioning in a Homogeneous Visual Field 405
Kennedy, A., and Galloway, A., Retrieval Time and the Effect of Serial Position 430
Kinney, J. A. S., Luria, S. M., and Weissman, S., Estimates of Size and Distance Underwater 282
Koh, S. D., Time-Error in Comparisons of Preferences for Musical Excerpts 171
Koplin, J. H., and Moates, D. R., Convergent Associations in Mediated Generalization 81
Krauskopf, J., Heterchromatic Stabilized Images: A Classroom Demonstration 634
Lamm, H., Coffman, T. L., Karlins, M., and Schroder, H. M., Conceptual Complexity, Recall, and Identification 631
Landauer, A. A., Blessing, W. W., and Coltheart, M., The Effect of False Perspective Clues on Distance- and Size-Judgments: An Examination of the Invariance Hypothesis 250
Leibowitz, H. W., and Judisch, J. M., The Relation between Age and the Magnitude of the Ponzo Illusion 105
Leibowitz, H. W., and Judisch, J. M., Size-Constancy in Older Persons: A Function of Distance 294
Leibowitz, H. W., Dickson, D., and Pollard, S. W., Monocular and Binocular Size-Matching as a Function of Distance at Various Age-Levels 263
Levy, B. A., Dockrill, F. J., and Mc Nulty, J. A., The Subthreshold Perception of Stimulus-Meaning 28
Lewis, T. C., and Jarrard, L. E., Effects of Hippocampal Ablation and Intertrial Interval on Acquisition and Extinction of a Complex Maze 66

LOEB, M., and GUNN, W. J., Correlations of Performance in Detecting Visual and Auditory Signals 236
LOIZOS, C. M., and ANSTIS, S. M., The Cross-Modal Judgment of Small Holes 51
LUDVIGSON, H. W., and CAUL, W. F., Spatial Generalization as a Function of Variation in the Nonspatial Attribute, Instructions, and Individual Strategies 186
LURIA, S. M., Color-Name as a Function of Stimulus-Intensity and Duration 14
LURIA, S. M., KINNEY, J. A. S., and WEISSMAN, S., Estimates of Size and Distance Underwater 282
MCALLISTER, D. E., and MCALLISTER, W. R., Drive and Reward in Aversive Learning 377
MCNULTY, J. A., DOCKRILL, F. J., and LEVY, B. A., The Subthreshold Perception of Stimulus-Meaning 28
MAIER, N. R. F., JULIUS, M., and THURBER, J. A., Studies in Creativity: Individual Differences in the Storing and Utilization of Information .. 492
MARKS, L. E., Judgments of Grammaticalness of Some English Sentences and Semi-Sentences 196
MARSHALL, A. J., and FITZGERALD, R. E., Left-Right Field Differences with Partial Report of Letters 370
MELTZER, L., and HAYES, D. P., Bone-Conducting Microphones 619
MICHON, J. A., The Game of *JAM*—An Isomorph of Tic-Tac-Toe 137
MITCHELL, D., DES LAURIERS, G., and WONG, R., Serial Learning of Children as Facilitated by Mediation 140
MOATES, D. R., and KOPLIN, J. H., Convergent Associations in Mediated Generaliation 81
MORGAN, A. H., COOPER, L. M., HILGARD, E. R., and VOEVODSKY, J., The Measurement of Suprathreshold Pain 124
NATSOULAS, T., Alternative Interpretation of Cutaneous Perception 313
NEWMAN, E. B., Seventy-Fourth Annual Convention of the American Psychological Association 144
NEWMAN, S. E., Effects of Serial Presentation of Paired Associates on Performance during and after Training 100
ODOM, P. B., and BLANTON, R. L., Rule Learning in Deaf and Hearing Subjects 391
PAPPAS, B. A., HEPBURN, K. L., and SUBOSKI, M. D., A Comparison of Recall and Binary Recognition 398
PICKENS, R. W., and CROWDER, W. F., A Recorder of Locomotor Activity .. 442
PLATT, J. R., and SPENCE, K. W., Intertrial Interval and Response-Competition in Discrete-Trials Lever-Pressing 110
POLLARD, S. W., DICKSON, D., and LEIBOWITZ, H. W., Monocular and Binocular Size-Matching as a Function of Distance at Various Age-Levels 263
PORSOLT, R. D., and IRWIN, R. J., Binaural Summation in Loudness of Two Tones as a Function of Their Bandwidth 384
POSTMAN, L., Sixty-Third Annual Meeting of the Society of Experimental Psychologists 465
RABBITT, P. M., Learning to Ignore Irrelevant Information 1
RAND, G., GLICK, J. A., WAPNER, S., and WEINBERG, J., Effect of Speed of Movement on Tactual Kinesthetic Perception of Extent 608
RAY, A. J., JR., Avoidance-Learning in the Tiger Salamander 642
RIMOLDI, H. J. A., and DONNELLY, M. B., Experimental Considerations Concerning Category and Magnitude Scaling 602
ROBY, T. B., and SINGER, E., Measurement of Sequential Behavior by Successive Differences 276
SARRIS, V., Adaptation-Level Theory: Two Critical Experiments on Helson's Weighted-Average Model 331
SARRIS, V., Comments on Helson and Masters' "Study of Inflection-Points in the Locus Adaptation-Level as a Function of Anchor-Stimuli" 304

SCHIFFMAN, H. R., Size-Estimation of Familiar Objects under Informati and Reduced Conditions of Viewing 229
SCHNALL, M., Global Errors in Children's Reproductions of Progressively Changing Pictures 213
SCHRODER, H. M., COFFMAN, T. L., KARLINS, M., and LAMM, H., Conceptual Complexity, Recall, and Identification 631
SEGAL, S. J., and GLICKMAN, M., Relaxation and the Perky Effect: The Influence of Body Position in Judgments of Imagery 257
SERSEN, F. A., VETTER, R. J., and WEINSTEIN, S., Preference for Bodily Parts Following Mastectomy 458
SHAW, W. A., and BECK, J., Ratio-Estimation of Loudness-Intervals 59
SHICKMAN, G. M., An Antilogarithmic Grid and Some Applications 128
SIEGEL, M. H., Stimulus-Repetition and Color Discrimination 411
SILVERSTEIN, A., The Prediction of Individual Association-Hierarchies from Cultural Frequencies 88
SINGER, E., Ability and the Use of Optimal Strategy in Decisions 243
SINGER, E., and ROBY, T. B., Measurement of Sequential Behavior by Successive Differences 276
SPENCE, K. W., and PLATT, J. R., Intertrial Interval and Response-Competition in Discrete-Trials Lever-Pressing 110
SPITZ, H. J., Ganz's Hypothesis of Figural After-Effects 462
STABLER, J. R., and TAYLOR, A., Performance in a Runway 95
STICHT, T. G., The Grayson-Stadler Electronic Switch in Reaction-Time Measurements 625
STONE, L. A., Illusory Motion 142
SUBOSKI, M.D., HEPBURN, K. L., and PAPPAS, B. A., A Comparison of Recall and Binary Recognition 398
TAYLOR, A., and STABLER, J. R., Performance in a Runway 95
THOMAS, J. P., Equipment for Varying the Intensity of Light 297
THURBER, J. A., JULIUS, M., and MAIER, N. R. F., Studies in Creativity: Individual Differences in the Storing and Utilization of Information .. 492
VETTER, R. J., SERSEN, F. A., and WEINSTEIN, S., Preferences for Bodily Parts Following Mastectomy 458
VOEVODSKY, J., COOPER, L. M., HILGARD, E. R., and MORGAN, A. H., The Measurement of Suprathreshold Pain 124
WADE, N. J., and DAY, R. H., Tilt and Centrifugation in Changing the Direction of Body-Force 637
WAPNER, S., GLICK, J. A., RAND, G., and WEINBERG, J., Effect of Speed of Movement in Tactual Kinesthetic Perception of Extent 608
WEBB, D., and GUMENIK, W. E., Effects of the Relevance of Interpolated Stimuli on Time-Error in the Comparative Judgment of Weights 115
WEINBERG, J., GLICK, J. A., RAND, G., and WAPNER, S., Effect of Speed of Movement in Tactual Kinesthetic Perception of Extent 608
WEINSTEIN, S., SERSEN, F. A., and VETTER, R. J., Preferences for Bodily Parts Following Mastectomy 458
WEISSMAN, S., KINNEY, J. A. S., and LURIA, S. M., Estimates of Size and Distance Underwater 282
WONG, R., DES LAURIERS, G., and MITCHELL, D., Serial Learning of Children as Facilitated by Mediation 140
WOODRUFF, B., and HELSON, H., Torque-Sensitivity as a Function of Knob-Radius and Load 558
WORCHEL, S., and BURNHAM, C. A., Reduction of Autokinesis with Information about the Registration of Eye-Position 434
YOUSSEF, Z. I., Association and Integration in Serial Learning 355

Book Reviews

(The reviewer's name appears in parentheses after the title of the work.)

ALPERT, R., *et al.*, Identification and Child Rearing (A. Stein) 481
APPLEY, M. H., and COFER, C. N., Motivation: Theory and Research (T. A. Ryan) 319
ATKINSON, J. W., An Introduction to Motivation (T. A. Ryan) 319
ATKINSON, R. C., *et al.*, An Introduction to Mathematical Learning (H. Bernbach) 482
BERITOFF, J. S., Neural Mechanics of Higher Vertebrate Behavior (R. L. Isaacson) 478
BOWER, G. H., *et al.*, An Introduction to Mathematical Learning (H. Bernbach) 482
BROWN, R., Social Psychology (A. Jenness) 324
BUSH, R. R., *et al.*, Handbook of Mathematical Psychology (R. McGinnis) 322
BUSH, R. R., *et al.*, Readings in Mathematical Psychology (R. McGinnis) 322
CAMPBELL, D. P., The Results of Counseling: Twenty-Five Years Later (A. G. Nelson) 483
CAMPBELL, D. T., *et al.*, Unobtrusive Measures: Nonreactive Research in the Social Sciences (P. Gump) 161
COFER, C. N., and APPLEY, M. H., Motivation: Theory and Research (T. A. Ryan) 319
CROTHERS, E. J., *et al.*, An Introduction to Mathematical Learning (H. Bernbach) 482
DEESE, J., The Structure of Associations in Languages and Thought (M. Anisfeld) 646
DOHRENWEND, B. S., *et al.*, Interviewing: Its Forms and Functions (J. H. Gustad) 484
EPSTEIN, W., Varieties of Perceptual Learning (J. F. Wohlwill) 474
EYSENCK, H. J., and RACHMAN, S., The Causes and Cures of Neurosis (N. D. Sundberg) 470
FOULKES, D., The Psychology of Sleep (D. Foster) 154
GAITO, J., Molecular Psychology: A Clinical Approach to Learning and Other Behavior (M. R. Rosenweig) 656
GALANTER, E., *et al.*, Handbook of Mathematical Psychology (R. McGinnis) 322
GALANTER, E., *et al.*, Readings in Mathematical Psychology (R. McGinnis) 322
GIBSON, J. J., The Senses Considered as Perceptual Systems (E. G. Boring) 150
GREGORY, R. L., Eye and Brain: The Psychology of Seeing (J. Beck) 159
GUION, R. M., Personnel Testing (P. C. Smith) 326
HALL, C. S., and VAN DE CASTLE, R. L., The Content Analysis of Dreams (W. Stephenson) 156
HALL, J. F., The Psychology of Learning (W. A. Russell) 650
HERING, E., Outlines of a Theory of the Light Sense (R. B. MacLeod) 163
HIMWICH, H. E., *et al.*, Amines and Schizophrenia (A. C. Goldstein) 659
HOWARD, I. P., and TEMPLETON, W. B., Human Spatial Orientation (H. L. Pick) 476
HUNT, F. R., *et al.*, Experiments in Induction (H. A. Simon) 651
HYMAN, R., The Nature of Psychological Inquiry (G. A. Ferguson) 658
KAHN, R. L., and KATZ, D., The Social Psychology of Organization (R. M. Stogdill) 480
KATZ, D., and KAHN, R. L., The Social Psychology of Organization (R. M. Stogdill) 480
KETTY, S. S., *et al.*, Amines and Schizophrenia (A. C. Goldstein) 659
KIRSCH, J., Shakespeare's Royal Self (L. J. Teagarden) 485
KLEIN, D., *et al.*, Interviewing: Its Forms and Functions (J. H. Gustad) 484
LOUCH, A. R., Explanation and Human Action (M. Black) 655
LUCE, R. D., *et al.*, Handbook of Mathematical Psychology (R. McGinnis) 322
LUCE, R. D., *et al.*, Readings in Mathematical Psychology (R. McGinnis) 322
MARIN, J., *et al.*, Experiments in Induction (H. A. Simon) 651

Murray, E. J., Sleep, Dreams, and Arousal (D. Foster) 154
Rachman, S., and Eysenck, H. J., The Causes and Cures of Neurosis (N. D. Sundberg) 470
Rau, L., *et al.*, Identification and Child Rearing (A. Stein) 481
Richardson, S. A., *et al.*, Interviewing: Its Forms and Functions (J. H. Gustad) 484
Ricks, D. F., and Wessman, A. E., Mood and Personality (V. Nowlin) .. 146
Schultz, D. P., Sensory Restriction: Effects on Behavior (D. W. Fiske) .. 661
Schwartz, R. D., *et al.*, Unobtrusive Measures: Nonreactive Research in the Social Sciences (P. Gump) 161
Sears, R. R., *et al.*, Identification and Child Rearing (A. Stein) 481
Sechrest, L., *et al.*, Unobtrusive Measures: Nonreactive Research in the Social Sciences (P. Gump) 161
Siegel, S., Choice, Strategy, and Utility (L. D. Phillips) 660
Smythies, J. R., *et al.*, Amines and Schizophrenia (A. C. Goldstein) 659
Stanisland, A. C., Patterns of Redundancy: A Psychological Study (H. W. Hake) 472
Stone, P., *et al.*, Experiments in Induction (H. A. Simon) 651
Templeton, W. B., and Howard, I. P., Human Spatial Orientation (H. L. Pick) 476
Thompson, R. F., Foundations of Physiological Psychology (B. H. Halpern) 653
Tinker, M. A., Bases for Effective Reading (J. Chall) 661
Van de Castle, R. L., and Hall, C. S., The Content Analysis of Dreams (W. Stephenson) 156
Veldman, D. J., Fortran Programming for the Behavioral Sciences (V. E. McGee) 657
Webb, E. J., *et al.*, Unobtrusive Measures: Nonreactive Research in the Social Sciences (P. Gump) 161
Wessman, A. E., and Ricks, D. F., Mood and Personality (V. Nowlin) .. 146

The American Journal of Psychology

The American Journal of Psychology

The American Journal of Psychology

THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. STANLEY HALL

VOL. LXXX MARCH, 1967 NO. 1

LEARNING TO IGNORE IRRELEVANT INFORMATION

By PATRICK M. RABBITT, Cambridge, England

In most tasks of choice-response, the time taken to discriminate between members of a set of stimuli varies with the size of the set, *i.e.* with the information-load. Exceptions to this rule occur in visual search.

For example, in tasks in which subjects (*S*s) searched for varying numbers of letters of the alphabet on displays containing all 26 letters. Shepherd,[1] and, in a different experimental situation, Rabbitt[2] found that the time taken to discriminate between relevant and irrelevant letters increased with the number of letters for which the *S*s searched, although the total vocabulary of relevant plus irrelevant letters was always the same (*i.e.* 26). It follows that the *S*s could not have scanned displays to identify each symbol in turn as a particular letter of the alphabet, responding appropriately only when they encountered a relevant letter. If they had, discrimination-time could only have varied with the total number of relevant plus irrelevant letters, and not with the ratio of relevant to irrelevant items, as was in fact observed. The results suggest that while the *S*s had to discriminate relevant from irrelevant letters they did not discriminate irrelevant items from other irrelevant items. The usual assumptions of information-theory must therefore be modified before they can be applied to visual search.

Rabbitt suggested that visual search might be treated as a particular example of a more general range of tasks in which the *S*s make the same response to all members of one group of stimuli and make different responses to members of other groups.[3] For example, in Shepherd's experiment the *S*s cancelled all relevant letters and ignored all irrelevant letters.[4]

Unfortunately, differences in the times taken to discriminate between various subsets of relevant and irrelevant items may be unreliable indices of the relative amounts of information which the *S*s have to transmit. There is now some doubt as to

---

* Received for publication April 25, 1966. From M. R. C., Applied Psychology Research Unit, Cambridge, England.

[1] R. D. Shepherd, Classification time and class complexity in inspection tasks, *Occ. Psychol.,* 38, 1964, 86–89.

[2] P. M. A. Rabbitt, Ignoring irrelevant information, *Brit. J. Psychol.,* 55, 1964, 403–414.

[3] Rabbitt, *op. cit.,* 411.

[4] Shepherd, *op. cit.,* 88.

whether any association between choice-response time and the number of S–R alternatives is observed after extended practice.[5] In the specific situation of visual search Neisser reports that the *Ss* scanning columns of symbols take no longer to locate a member of a set of 10 different relevant items than to locate a single relevant item.[6] It is doubtful whether the discrepancy between Neisser's results, on the one hand, and Rabbitt's and Shepherd's findings, on the other, can be reconciled by experiments based on comparisons between response-times obtained in different situations, since all such comparisons can be criticised on the grounds that they do not represent some hypothetical limiting level of performance attained after extended practice.[7]

Letters are complex stimuli which may be discriminated from each other in terms of the presence or absence of a wide range of characteristics. Fitts *et al.* have shown that *Ss* learning to discriminate between complex patterns improve their performance by identifying and using a range of critical cues and by ignoring other details as misleading or redundant.[8] The process of discriminating relevant from irrelevant letters would clearly be most economical if the *Ss* used the minimal number of cues. This minimal number might be expected to vary with the ratio of relevant to irrelevant items rather than directly with the total vocabulary of items on a display. We might also expect that the number of cues which *S* must consider (and so, possibly, also his decision-time) would increase with the number of items in the relevant set.

It has been suggested that improvement in speed of visual search with practice depends partly on the learning of a cue-system which is optimal for discriminations between specific sets of relevant and irrelevant items.[9] If this is so, substitution of a new set of irrelevant items would slow search. In this case the slowing of search on transfer from one display of irrelevant symbols to another will be an index of the extent to which previous improvement with practice has been due to learning of a specific cue-system.

Experiment I: Two or Eight Relevant Stimuli

This index is used in Experiment I to compare the relation of improvement with practice to learning of a specific, relevant-irrelevant cue-system as the number of relevant stimuli is increased from two to eight.

Material. Two packs of 48 cards measuring 2¼ × 3½ in. were made up for each of two conditions:

Condition 1. Half the cards in each of these two packs were stencilled, in India ink, with the letter A and half with the letter B. Letters were block-capitals 6 mm. high. These were the relevant letters for which the *Ss* searched, sorting the packs to separate A and B cards. The relevant letter on each card was imbedded in a 9-letter nonsense-word centered on the card. The other (irrelevant) letters making up this word were eight letters of the alphabet stencilled in the same ink, size and style as the relevant letter. All cards within each of these two packs bore the same irrelevant letters. In Pack 1 these were R, V, N, Q, Z, O, T, and J. In Pack 2 they were Z, M, S, I, U, P, W, and L. The positions of the irrelevant letters were so varied that every card bore a different nonsense-word. The position of the relevant letter was also systematically varied so that, as far as possible, A and B appeared equally often in each of the nine possible letter-positions.

---

[5] E. H. Mowbray and M. V. Rhoades, On the reduction of choice-response times with practice, *Quart. J. exp. Psychol.*, 11, 1959, 16–24.

[6] Ulric Neisser, Decision-times without reaction-time: Experiments in visual scanning, JOURNAL, 76, 1963, 376–385.

[7] Neisser, Visual search, *Scient. Amer.*, 210, 1964, 94–102.

[8] P. M. Fitts, Meyer Weinstein, Maurice Rappaport, Nancy Anderson and J. A. Leonard, Stimulus correlates of visual pattern-recognition, *J. exp. Psychol.*, 51, 1956, 1–11.

[9] Rabbitt, *op. cit.*, 410 f.

Condition 2. Six cards in each pack were stencilled with one of the eight relevant letters, A, B, C, D, E, F, G, and H. *Ss* sorted these packs into two piles to separate cards marked with A, B, C, or D from cards marked with E, F, G, or H. The relevant letter on each card was embedded in a 9-letter nonsense-word as in Condition 1. The two packs featured the same sets of irrelevant letters used in Condition 1. The same precautions were taken to randomize the positions of relevant and irrelevant letters within each pack.

Procedure. Each *S* was given one of the two packs in his condition and asked to sort it into two piles on the scheme described. He might be required to sort the same pack 1, 3, 6, 9, or 12 times, depending on the group to which he was assigned. After these training trials he was given the other pack in the same condition, without comment on the exchange, and asked to sort it as fast as possible.

Five groups of 18 *Ss*, in each of the conditions, each experienced one of the 5 levels of practice described (*i.e.* 1, 3, 6, 9, or 12 practice sorts before transfer). *E* thoroughly shuffled each pack before it was sorted and timed every sort with a stop-watch. The *Ss* were urged to sort as fast and accurately as possible; improvement was praised and errors were recorded and admonished. Packs of cards were held face-down for sorting, cards being turned over one at a time for inspection to prevent preview of sequent displays. Half the *Ss* in each group of 18 were trained on one pack in the condition while the remainder reversed this order of practice. All *Ss* were tested individually.

Subjects. One hundred and eighty *Ss* were tested. Ninety were unpaid volunteer Royal Navy "ratings" aged 17–31 yr. (mean 21.7), and 90 were paid housewives aged 22–48 yr. (mean age 39.2).

The *Ss* were divided into 10 groups of 18 (5 groups per condition) with 9 'ratings' and 9 housewives in each group. The 'ratings' form an unusually homogeneous population and a random allocation resulted in groups closely matched for age and *I.Q.* The housewives were also allocated at random, with precautions to match groups as closely as possible for age.

All the *Ss* were free of gross visual defects and wore any correcting lenses prescribed for them.

Results: (1) Errors. No groups in either condition made errors on more than 0.3% of responses. This incidence was too small to allow valid comparisons.

(2) Sorting-time. Mean sorting-times are set out in Table I.

TABLE I

MEAN SORTING-TIMES PER PACK ON TRAINING- AND TRANSFER-TRIALS

Groups		Practice trials: 1	2	3	4	5	6	7	8	9	10	11	12	Transfer
Searching for 2 relevant items	$\bar{X}$ sec.	91.0	81.2	77.3	76.4	71.1	73.9	73.4	72.8	70.1	68.9	67.3	68.5	80.1
	σ	18.1	20.2	17.4	18.1	15.2	14.3	13.3	18.2	12.4	14.2	12.4	12.2	14.3
	$\bar{X}$	89.0	80.4	78.6	75.6	72.7	72.8	72.2	71.1	68.4				79.1
	σ	20.5	21.4	27.6	16.9	15.9	16.7	15.6	16.2	18.2				15.4
	$\bar{X}$	78.0	78.1	78.2	74.5	71.8	71.9							80.9
	σ	19.7	16.9	15.4	13.2	15.4	13.4							14.6
	$\bar{X}$	93.0	80.8	78.7										83.2
	σ	14.9	17.2	14.9										15.2
	$\bar{X}$	86.0												83.7
	σ	14.1												14.7
Searching for 8 relevant items	$\bar{X}$	150.9	133.0	116.4	114.0	109.7	107.2	98.4	98.5	95.2	96.4	95.8	95.4	114.9
	σ	34.4	30.2	24.1	20.1	23.4	22.1	23.4	18.2	18.0	17.1	17.4	17.0	20.2
	$\bar{X}$	152.6	137.5	118.3	115.3	107.6	105.4	97.9	98.1	94.0				113.5
	σ	31.4	29.3	28.3	28.6	21.2	23.4	19.6	22.0	18.8				23.8
	$\bar{X}$	151.8	131.2	117.6	111.5	107.1	99.8							114.0
	σ	31.7	25.4	24.2	24.0	19.6	17.6							19.7
	$\bar{X}$	154.3	132.0	119.3										116.0
	σ	41.6	29.5	20.5										19.6
	$\bar{X}$	149.8												131.4
	σ	21.4												19.4

Homogeneity of populations. Comparison between groups of *Ss* brought to different levels of practice within each condition involves the assumption that the groups are appropriately matched. To test this, analyses of variance were made on scores for the first practice trial over all five groups, over the first three trials for the four groups, over the first six practice-trials for three groups, and over the first nine practice trials for two groups. None of these analyses showed any significant difference between groups. While no similar check could be made between the *Ss* in Conditions 1 and 2, the large numbers involved, and the consistency within each sample, make homogeneity a reasonable assumption.

(3) Decrement on transfer to a different irrelevant set. Mean sorting times for the last practice and the transfer-trial were compared by *t*-test at each level of practice within each condition. In Condition 1 (2-relevant), sorting-times on the transfer trial were not significantly different from sorting-times for the first trial on the practice-pack ($p > 0.1$). Times for transfer-packs were significantly longer, however, than for the terminal trial when this was the 3rd ($p < 0.01$), the 6th ($p < 0.001$), the 9th ($p < 0.001$) or the 12th ($p < 0.001$) practice-trial. In Conditions 2 (8 relevant), the mean sorting time on transfer was significantly *faster* after 1 practice-trial ($p < 0.01$), and there was no significant difference after 3 practice-trials ($p > 0.1$). Sorting-times for transfer-packs were significantly longer for the final trial after 6 ($p < 0.01$) after 9 ($p < 0.001$) and after 12 ($p < 0.001$) practice-trials.

Mean difference-scores (terminal practice–transfer-trial) for each group are plotted in Fig. 1 as a function of practice in both conditions. Negative values indicate that the transfer-packs were sorted faster than the practice-packs.

The S^2 statistic (studentized range statistics), was calculated for the mean difference scores over all groups.[10] In Condition 1 (2-relevant), mean decrement on transfer was significantly greater after 3 training-trials than after 1 training-trial ($p < 0.01$), and greater after 6 training-trials than after 3 training-trials ($p < 0.01$). The decrement differences between 9 and 12 practice-trials were not significant ($p > 0.1$). In Condition 2 (8-relevant) differences between 1, 3, 6, and 9 practice-trials were all significant at the ($p < 0.01$) level. The difference-scores for 9 and 12 practice-trials were not significant. Using S^2 it was also possible to compare decrement-scores between conditions at equivalent levels of practice. After one practice trial, decrement in the 2-relevant condition was significantly greater than in the 8-relevant condition ($p < 0.01$). There was no difference between conditions after three trials ($p. > 0.25$), and decrement scores were significantly greater in the 8-relevant condition that in the

---

[10] B. J. Wimer, *Statistical principles in experimental design*, 1962, 72.

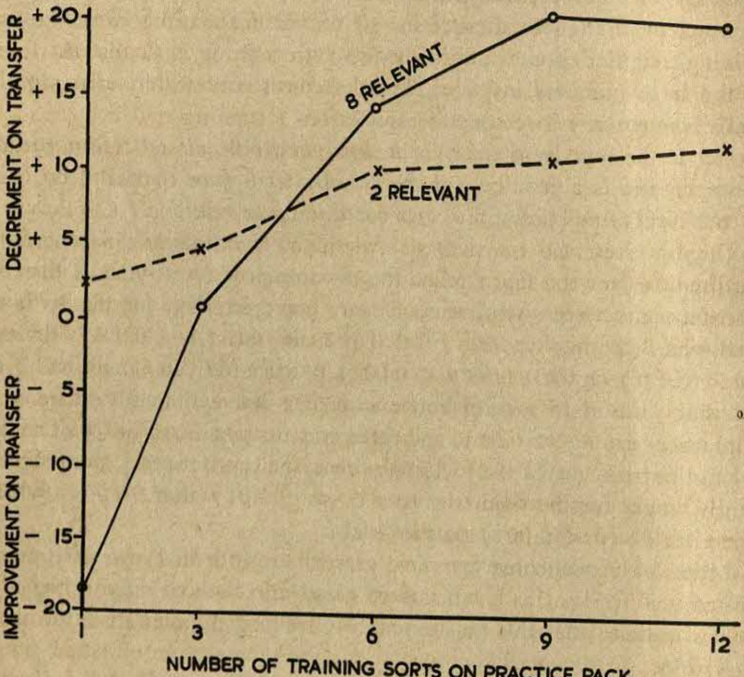


FIG. 1. MEAN DIFFERENCE-SCORE, TRANSFER-TRIAL, LAST PRACTICE TRIAL (SEC.)

2-relevant condition after six ($p < 0.01$) nine ($p < 0.001$) and twelve trials ($p < 0.001$).

Discussion. Decrements in speed on transfer show that improvement with practice is partly due to specific learning of a particular set of irrelevant letters. It is reasonable to use these decrements as indices to compare the improvements due to such specific learning for different levels of practice and conditions. Although curves for this specific learning cannot be directly obtained, the curves showing decrements in transfer as a function of practice in each condition (Fig. 1) may be considered analogs of these functions.

On this analogy, we see from Fig. 1 that *S*s searching for only two relevant symbols show rapid specific learning between one and six training trials. At this point the curve appears to reach an asymptote and there is no suggestion that decrement increases with further practice. In contrast, *S*s searching for 8 relevant items show continuing learning until the 9th training trial, and it is not until the 9th-12th training trials that there are signs that learning is completed. The decrement on transfer varies significantly, moreover, between conditions. After 1 training trial decrement is greater in the 2-relevant than in the 8-relevant task, after 3 training trials decrement is the same in both the tasks but after 6, 9 or 12 trials decrement on transfer is uniformly greater in the 8-relevant task.

The data thus fit the hypothesis that the *S*s have to learn more cues to discriminate between eight relevant and eight irrelevant symbols than to discriminate between two relevant and eight irrelevant symbols. It is consistent with this hypothesis that if the cue-system required to discriminate eight relevant letters is more elaborate it should take longer to learn. Again, because it is more elaborate, once it is learned, its loss on transfer might well imply correspondingly greater reduction in efficiency. The results of Experiment I therefore fit the general hypothesis that the difficulty of discriminations between two sets of items varies with the ratio in which items are distributed between the sets.

Letters of the alphabet are complex symbols and the range of cues required to discriminate between them may vary between wide upper and lower limits which can only be roughly fixed by the number of symbols in each set. Although Experiment I suggests that the information transmitted in discriminating between sets of symbols does, in general, vary with the number of symbols to be discriminated, nevertheless the number of necessary cues may vary with the particular physical characteristics of different sub-sets. Variations in the number of items in the sets to be discriminated may prove only roughly correlated with changes in the amount of information which the subject transmits.

EXPERIMENT II: VARIATIONS IN CUE-SYSTEM

Experiment II was made to provide a practical demonstration of this point as a basis for further discussion, and also as a further test of the sensitivity of the technique used in Experiment I.

Material. Three packs of 50 cards were made up, as described for Experiment I. Half the cards in each pack each bore the relevant letter C and the other half the relevant letter O. In all packs, these relevant letters were embedded, among sets of seven irrelevant letters forming with them an 8-letter nonsense-word centered on each card. In Pack 1, irrelevant letters contained no curved lines (A, E, F, H, I, K, and L). The same was true of Pack 2, in which the irrelevant letters were M, N, T, V, W, X, and Y. In Pack 3 the rounded relevant letters C and O were embedded among other letters contained curved lines (B, D, G, J, P, Q, and S). The same precautions described for Experiment I were taken to randomize the locations of relevant and irrelevant letters on the displays.

Procedure. Two groups of 24 *Ss* were tested. Half the *Ss* in each group sorted Pack 1 six times, the remainder sorted Pack 2 six times. After this initial training, all the *Ss* in Group 1 were transferred to Pack 3. In Group 2, the *Ss* practiced on Pack 1 were transferred to Pack 2 and vice-versa. This tested the extent to which training on discriminations between rounded and angular letters transferred when a new set of angular letters was used.

The procedure for administering and timing the packs was the same as in Experiment I.

Subjects. Twelve *Ss* in each group were R.N. ratings aged from 17–24 yr. (mean age 19.8). The remaining 12 *Ss* in each group were housewives aged from 32–46 yr. (mean age 38.2). Groups were matched for age as closely as possible.

TABLE II
MEAN SORTING-TIMES PER PACK ON TRAINING AND TRANSFER TRIALS

		Transfer to pack 1	Transfer to pack 2	Transfer to pack 3
Pack 1	66.2 ($\sigma=4.9$)		67.1 ($\sigma=5.4$)	79.4 ($\sigma=8.2$)
Pack 2	65.1 ($\sigma=4.2$)	66.4 ($\sigma=5.0$)		78.4 ($\sigma=9.0$)

Results. Mean sorting times for each group on each practice trial and on the transfer-trial are set out in Table II.

Sorting times on the six practice trials are closely similar for each group, and analysis of variance shows no significant difference between them. The *Ss* in Group 1 show marked decrement on transfer from discriminations between rounded and angular letters to discriminations between rounded letters and other rounded letters ($p < 0.001$). In contrast, the *Ss* in Group 2 show no significant decrement on transfer to a new set of irrelevant letters ($p > 0.1$). We must conclude that while the same cue system is used to discriminate round from angular letters in Packs 1 and 2, different cue-systems are required to discriminate relevant from irrelevant letters in Packs 1 and 2 on the one hand and in Pack 3 on the other.

Discussion. As far as we can deduce from Experiment II, the same cue-system was used to discriminate the letters C and O from either of the two sets of angular irrelevant letters, *i.e.* the same cue-system was appropriate for both sub-sets of 7 out of 14 angular letters. Thus, under these circumstances, variation in the vocabulary of irrelevant letters is seen to be a less critical determinant of choice-time than variation in the physical characteristics of the particular relevant and irrelevant items.

The results of Experiments I and II throw some light on discrepancies in the results of earlier investigations on visual search; Shepherd and Rabbitt found that the time taken to discriminate relevant from irrelevant stimuli on a display increased with the number of relevant stimuli for which their *S*s searched.[11] In contrast, Neisser reports that the *S*s scanning columns of symbols take no longer to locate any one of a set of 10 different items than to locate a single relevant item.[12]

The task of discriminating all members of a set of relevant items from all members of a set of irrelevant items is, logically, a particular example of the general situation in which the *S*s must make one response to all members of one set of stimuli and another response to all members of another similar set. It is not therefore surprising that discrepancies in results with visual search should be paralleled in the literature on stimulus-categorization tasks. Nickerson and Feehrer found that the time taken to identify stimuli as members of one of two categories increased as a linear function of the number of stimuli in each category.[13] In another task of this kind, however, Pollack reports that while choice time rises sharply as the number of stimuli in each of 2 categories is increased from 1 to 2, further increases in stimulus-entropy have markedly less effect.[14] Rabbitt also found that *S*s choosing between 2 classes of stimuli showed a sharp increase in response-time as the number of stimuli in each class increased from 1 to 2 items, but that further increases from 2 to 12 items resulted in no detectable change in response-time.[15] Finally, in contrast to all these other results, Broadbent and Gregory found no change in response-time as the number of stimuli in each of 4 response-classes was increased from 1 to 8 items.[16]

Comparisons between these four sets of data are confounded by important differences in the tasks from which they were obtained. Nevertheless, it is probable that an important variable was the amount of practice which the *Ss* received. For example Nickerson and Feehrer gave their *Ss* less practice than did any other investigator quoted,[17] while Broadbent and Gregory investigated a task in which the *Ss* were aided by their life-experience in reading.[18] Broadbent and Gregory's *Ss* separated cards marked with the name of one British town from cards marked with the names of others. Stimulus-entropy was varied within each class by printing these town names in different combinations of upper and lower case letters. Thus the speed with which *S* identified, say, the combination bOlTOn and BoltOn as members of the same class related to a lifetime's experience at reading meaningful combinations of letters. A training study by Rabbitt directly makes the point that some interactions between stimulus-entropy and response-time may disappear with extended practice.[19]

---

[11] Rabbitt, *op. cit.*, 406 ff.; Shepherd, *op. cit.*, 87–97.

[12] Neisser, *op. cit.*, *Scient. Amer.*, 94–102.

[13] R. S. Nickerson and C. E. Feehrer, Stimulus categorisation and response-time, *Percep. mot. Skills*, 18, 1964, 785–793.

[14] Irving Pollack, Speed of classification of words into superordinate categories, *J. verb. Learn, verb. Behav.*, 2, 1963, 159–165.

[15] Rabbitt, Effects of independent variations in stimulus and response probability, *Nature*, 183, 1959, 1212.

[16] D. E. Broadbent and Margaret Gregory, Human responses to classes of stimuli, *Nature*, 192, 1962, 1315 f.

The analogy between search and categorization suggests that discrepancies between Neisser's, Shepherd's and Rabbitt's results may also be due to differences in practice.[20] This explanation must be treated with caution, because Shepherd's and Rabbitt's *Ss* were highly trained in their tasks. There are, moreover, wide differences in the tasks investigated by the three authors, and at least two general reasons why we cannot be sure that Neisser's results described a limiting level of performance which can be attained in *all* tasks of visual search.

(1) Experiment I shows that improvement with practice requires the learning of a system of cues which is optimal for discriminations between a particular set of relevant and a particular set of irrelevant symbols. It is also apparent that though the *Ss* may, with practice, reduce the range of critical cues which they consider in order to make their discriminations, the number of cues which they must use is likely to increase with the number of relevant symbols for which they search. Experiment II suggests that the vocabulary of *cues* that the *Ss* use may be independent of very wide variations in the vocabulary of irrelevant symbols amongst which they search. Neither Neisser nor Broadbent and Gregory suggests that the time taken to distinguish between any two or four classes of symbols is independent of *all* possible variations in the range of cues which *Ss* must consider —whether such variations are associated with the number or with the *complexity* of symbols to be discriminated. A similar point has been made by Conrad discussing his finding that short-term memory capacity for letters

---

[17] Nickerson and Feehrer, *op. cit.*, 785–793.

[18] Broadbent and Gregory, *op. cit.*, 1315 f.

[19] Rabbitt, Perceptual discrimination and the choice of responses. Unpublished Doctoral dissertation, University of Cambridge, 1962, 2.47 f.

[20] Niesser, *op. cit.*, this JOURNAL, 376–385; Shepherd, *op. cit.*, 87–97; Rabbitt, *op. cit.*, 403–414.

of the alphabet is limited by their acoustic confusability rather than by their information load as conventionally calculated.[21]

(2) Neisser's and Shepherd's tasks represent special cases of visual search in which their *S*s made the same response to all relevant stimuli. In contrast, *S*s in Experiment I above and in earlier studies made one response to all stimuli in one relevant sub-set (*e.g.* the letters A, B, C, or D) and another response to all stimuli in a different relevant sub-set (*e.g.*, the letters E, F, G, or H). This implies that they discriminated between three rather than between two categories of stimuli (Relevant$_1$/Relevant$_2$/ Irrelevant).

The results of previous experiments in which the *S*s have discriminated between more than two categories of stimuli suggest that response-time is a multiplicative function of the number of categories and of the number of stimuli in each such category. For example, an increase in the total stimulus ensemble from 12 to 16 items may cause little or no increase in response time when the *S*s discriminate them into four response-classes, but may bring about considerable increases in response-time when they discriminate them into six or more classes.[22] In visual search we might therefore also expect that the time taken to distinguish relevant from irrelevant stimuli would vary with the number of different categories into which the set of relevant stimuli must be separated.

It is obviously important to find some means of comparing the (various) amounts of information transmitted by the *S*s undertaking different search- or categorization-tasks. Unfortunately there are at least six objections to the conventional use of variation of choice-response time as an index for this purpose:

(a) Functions predicting changes in rate of gain of information must be based on comparisons among a wide range of situations in which populations of stimuli are divided into different numbers of equal and unequal sub-sets. The available data suggest that discrimination-times depend on some interaction between the level of practice, the number of stimuli involved, the number of sub-sets into which stimuli must be divided and the relative sizes of these different sub-sets. Empirical determinations of rates of gain of information in different categorizations therefore require the slippery decision as to when a "limiting level of performance after extended practice" is attained for each of a large number of different tasks.

(b) The relative sizes of the vocabularies of items between which the *S*s discriminate may be a very coarse index of vocabularies of *cues* which they use, and consequently also of the information which they transmit. (Experiment II above). The features employed as cues must vary with practice, with the number of items

---

[21] R. Conrad, Acoustic confusions in immediate memory, *Brit. J. Psychol.*, 55, 1964, 75–84.

[22] Pollack, *op. cit.*, 159–165; Rabbitt, *op. cit.*, *Nature*, 183, 1959, 1212 f.

in the particular sub-sets of items involved and, among other factors, with the relative perceptual "obviousness" of the various different available cue-systems.[23]

While it is indifferent to the calculation of transmitted information what particular features of complex stimuli are employed as cues, variations in the discriminability of different sets of potential cues must confound the use of choice-time as an index of the relative amounts of information transmitted in different situations. (See also Conrad for a discussion of this point in relation to the limits set to short-term memory capacity by confusability of items stored.)[24]

(c) It may be theoretically possible to calculate how the optimal range of cues needed to discriminate between given sets of stimuli will vary as the principles of categorization are altered. It is unlikely, however, that *S*s use *optimal* sets of cues. We only know that they reduce the range of cues they use as far as they are able.

(d) The existing data do not suggest that the time taken to discriminate between categories of items varies continuously with the range of cues which the *S*s must consider. The majority of experiments on stimulus categorization suggest some discontinuities in the relationship between stimulus-information and choice response-time.[25]

(e) As an extension of *(d)*, information transmitted in categorizing complex stimuli depends not on the principles of classification *required* of the *S* but rather on the strategy of classification which he actually *uses*. For example, a theoretically optimum strategy to discriminate the 8 letters, A, B, C, D, E, F, G, H, into the 4 sets A,B/C,D/E,F/G,H would use only such cues as are necessary to distinguish letters within each set from letters in other sets (*i.e.* the letter A need only be distinguished from the letters C, D, E, F, G, H but not from the letter B and so on). Cues which distinguish letters from others in the same set are theoretically redundant, and need not be considered. The *S*s may, however, perform this task in a sub-optimal way, and identify each letter as a particular one of a set of eight possible items and then assign it to its response-category. We know that *S*s do not adopt this sub-optimal strategy in some visual tasks, but we do not yet know the limiting range of cues within which such economical strategies are possible.

These five arguments also emphasize that comparisons of performance under various principles of stimulus-categorization (Broadbent and Gregory, Pollack, Nickerson, Feehrer & Rabbitt) provide very unreliable indices of the relative amounts of information processed in different situations.[26] An additional difficulty is specific to conditions encountered in visual search experiments:

(f) Where many symbols are simultaneously present on a display, *S*s are known to process them in groups of two or more spatially-adjacent items rather than one at a time.[27] Rabbitt has suggested that details of differences in the time taken to

---

[23] Rabbitt, Stimulus information load and discriminability, *Nature*, 197, 1963, 726 f.

[24] Conrad, *op. cit.*, 75–84.

[25] Broadbent and Gregory, *op. cit.*, *Nature*, 192, 1962, 1315 f., Neisser, *op. cit.*, this JOURNAL, 76, 1963, 376–385; Pollack, *op. cit.*, *J. verb. Learn. verb. Behav.*, 2, 1963, 159–165; Rabbitt, *op. cit.*, *Nature*, 183, 1959, 1212.

[26] Broadbent and Gregory, *op. cit.*, 1315; Pollack, *op. cit.*, 159 ff. Nickerson and Feehrer, *op. cit.*, 785 ff.; Rabbitt, *op. cit.*, *Nature*, 1212 f.

[27] N. H. Mackworth and J. F. Mackworth, Eye fixations recorded on changing visual scenes by the television eye-marker, *J. opt. Soc. Amer.* 48, 1958, 439–445.

search for two and for eight relevant items can only be explained on the assumption that *S*s can process several adjacent symbols together when searching for two items, but may not be able to do so when searching for as many as eight.[28] Similarly, Morton has shown that eye-voice span in reading (and also the number of symbols on a printed page considered at a single fixation) varies inversely with the information-load of the material read.[29] The trade-off between eye-voice span and information load which masks differences in reading speed with various materials may also be expected to operate in visual search.

The experiments described offer examples of a technique which may be useful in the investigation of the cue-systems which *S*s recognize and adopt in visual search tasks. The development of similar techniques must become increasingly important, since experiments to investigate perceptual discriminations between complex stimuli are of very limited value unless they are designed to reveal which of the very complex series of options available to them *S*s actually learn and use.

SUMMARY

Two experiments used a training-transfer paradigm to investigate learning of irrelevant symbols in visual search tasks. Groups of *S*s were practiced at locating members of particular sets of relevant letters embedded among restricted vocabularies of irrelevant letters. Search-time reduced with practice, but increased when *S*s were transfered to displays incorporating a new irrelevant vocabulary. This increase in search-time on transfer provided an index of the extent to which improvement related to learning of cues was specific to discriminations between particular relevant and irrelevant sets.

The results of Experiment I suggest that the number of cues used to discriminate relevant from irrelevant letters increases with the number of relevant letters for which the *S* is required to search. Experiment II suggested that search time is less affected by the relative sizes of the relevant and irrelevant vocabularies than by the particular physical characteristics of symbols in the relevant and irrelevant sets. These results are discussed in the context of recent experiments investigating the time taken to discriminate between classes of signals.

---

[28] Rabbitt, *op. cit., Brit. J. Psychol.,* 55, 1964, 406 ff.

[29] John Morton, The effects of context upon speed of reading, eye-movements, and eye-voice span, *Quart. J. exp. Psychol.,* 16, 1964, 340–354.

COLOR-NAME AS A FUNCTION OF STIMULUS-INTENSITY AND DURATION

By S. M. LURIA, U. S. Naval Submarine Medical Center, Groton*

Much of the previous work on color-naming has been aimed at determining the number of different colors which can be absolutely identified,[1] although Chapanis has gone into the problem more deeply, studying the number of names required to describe the total color-space, the denotative meaning of color-names, and related problems.[2]

In general, when color-names have been used as a dependent variable, the independent variables typically have been states of *O*. There is not much information on the effects of changes in stimulus-conditions on the color-names associated with various wavelengths.[3] Changes in such stimulus-conditions as intensity and duration produce, of course, marked changes in the color-matches made by the usual psychophysical methods.[4] In view of the discrepancy between the immense number of colors which can be discriminated and the small number of color-names which can be used effectively, it is of interest to see to what extent these stimulus-changes affect color-names. Thomson has reported some data,[5] and Boynton and Gordon have studied color-names at two luminance-levels.[6] A number of authors have investigated certain theoretical issues involving color-names, but generally under a restricted range of conditions.[7] Walraven has explored the entire spectrum under several con-

---

* Received for publication August 25, 1966, from Bureau of Medicine and Surgery, Navy Department, Research Work Unit MF022.03.03-9006.

[1] M. R. Jones, Color-coding, *Hum. Fac.*, 4, 1962, 355–365.

[2] Alphonse Chapanis, Color names for color space, *Amer. Sci.*, 53, 1965, 327–346.

[3] For studies of surface-colors see, for example, W. E. K. Middleton and E. G. Mayo, The appearance of colors in twilight, *J. opt. Soc. Amer.*, 42, 1952, 116–121; H. R. Davidson, Visual sensitivity to surface color differences, *ibid.*, 41, 1951, 104-111.

[4] Cf, for example, R. W. Burnham, The dependence of color upon area, this JOURNAL, 64, 1951, 521–533; D. Jameson and L. M. Hurvich, Effect of exposure-time on perceived color and color contrast, *J. opt. Soc. Amer.*, 52, 1962, 1326(A); J. A. S. Kinney, Changes in appearance of colored stimuli with exposure duration, *ibid.*, 55, 1965, 738–739.

[5] L. C. Thomson, Sensations aroused by monochromatic stimuli and their prediction, *Opt. Acta*, 2, 1954, 92–102.

[6] R. M. Boynton and J. Gordon, Bezold-Brücke shift measured by color-naming technique, *J. opt. Soc. Amer.*, 55, 1965, 78–86.

[7] F. L. Dimmick and M. R. Hubbard, The spectral location of psychologically unique yellow, green, and blue, this JOURNAL, 52, 1939, 242–254; The spectral components of psychologically unique red, this JOURNAL, 52, 1939, 348–353; John Krauskopf and R. Srebre, Spectral sensitivity of color mechanisms: derivation from fluctuations of color appearance near threshold, *Science*, 146, 1965, 1477–1479; John Krauskopf, Color appearance of small stimuli and the spatial distribution of color receptors, *J. opt. Soc. Amer.*, 54, 1964, 1171; Lawrence Wheeler,

ditions, but his interest has centered on quantum theory and the probabilities of changes near threshold.[8] Beare investigated two intensity-levels, 3000 and 7500 ml., and found only minor changes, but these intensities are rather high and differ by less than half a log unit.[9] There are indications, in addition to Boynton and Gordon's data, that lower intensities and short exposure-times would produce larger changes; for example, Rouse[10] and Luria and Weissman[11] found that it was difficult to distinguish green from blue with brief, dim stimuli.

It is the purpose of this study to determine the color-names associated with wavelengths whose intensity and duration varied through ranges of three log units. The color-naming technique developed by Boynton, Schafer and Neun[12] was used, a technique which Boynton and Gordon found to be "just as sensitive a method as is direct matching."[13]

Method: (1) Apparatus. A projection-system and sets of interference- and blocking filters were used rather than a monochromator to achieve the intensities desired. The spectrophotometric curves of the eleven sets of filters are shown in Fig. 1.

O sat with his head in a chin-rest before a white bainbridge-board screen which subtended a 70 × 100° visual angle. The center of the screen was illuminated by a 300-w. projection-lamp plus a Macbeth 'daylight' glass filter. Its luminance was controlled by layers of black cheesecloth, which makes a perfectly neutral filter. Additional illumination on the outer parts of the screen was provided by fluorescent bulbs. In the center of the screen, at a distance of 60 cm., was a circular hole of 45′ visual angle. In back of this hole were the filters, through which *O* looked directly at the 750-w. light-source. The luminance of the stimuli was controlled by neutral-density filters plus minor adjustments of the amperage of the light-source. The exposure time was controlled by a sectored disk and a constant-speed motor-and-clutch assembly positioned behind the filters.

(2) Method. The *O*s were instructed to use five color-names—*red, yellow, green, blue,* and *white*—and to identify the stimuli by using either one of the names or a combination of any two, the first color mentioned being the dominant one. Each response was worth three points. If only one name was used, it was assigned the

---

Color-naming responses to red light of varying luminance and purity, *J. opt. Soc. Amer.* 52, 1962, 1058–1066; M. A. Bouman and P. L. Walraven, Some colour naming experiments for red and green monochromatic lights, *J. opt. Soc. Amer.,* 47, 1957, 834–839; F. H. C. Marriott, Color naming experiments and the two-quanta theory, *J. opt. Soc. Amer.,* 49, 1959, 1022.

[8] P. L. Walraven, *On the Mechanisms of Colour Vision,* Institute for Perception RVO-TNO, Soesterberg, The Netherlands, 1962, pp. 47–59.

[9] A. C. Beare, Color-name as function of wave-length, this JOURNAL, 76, 1963, 248–256.

[10] R. O. Rouse, Color and the intensity-time relation, *J. opt. Soc. Amer.,* 42, 1952, 626–630.

[11] S. M. Luria and S. Weissman, Effect of stimulus duration on the perception of red-green and yellow-blue mixtures, *J. opt. Soc. Amer.,* 55, 1965, 1068–1072.

[12] R. M. Boynton, W. Schafer, and M. E. Neun, Hue-wavelength relation measured by color-naming method for three retinal locations, *Science,* 145, 1964, 666–668.

[13] Boynton and Gordon, *op. cit.,* 86.

full three points; if two names were used, the first was scored two points and the second one point.

(3) Procedure. Three luminance-levels on the surround (15, 1.9, and 0.15 ft-L.) and three exposure-times (300, 20, and 2 m.sec.) were used. Prior to the experiment, the brightness of each stimulus at each exposure-time was matched to each luminance-level of the surround. Since the equal energy stimuli appear dimmer at shorter durations, this was compensated for by increasing the energy of the briefer stimuli to keep brightness constant. *O*s were adapted to the lowest luminance-level for 4 min. All observations were monocular.

The various conditions were presented in a different random order to each *O*. Two conditions were presented during each session, and five judgments were made of each stimulus under each condition. Each condition was given three times to two *O*s and twice to the third *O*. Stimuli were presented at the rate of about one every 10 sec. Each session lasted about 40 min.

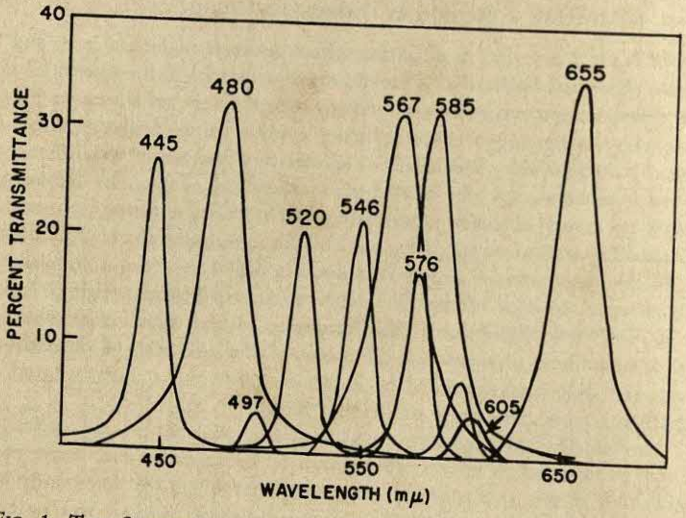


FIG. 1. THE SPECTROPHOTOMETRIC CURVES OF THE ELEVEN SETS OF STIMULUS-FILTERS

(4) Observers. Three color-normal *O*s and one deuteranope were tested. All were experienced in psychophysical observation. The deuteranope was tested for three sessions under each condition, but a total of only 10 judgments were made for each stimulus.

Results. There was not much variability between sessions for a given *O* and no marked differences between the normal *O*s except at 497 mμ. The scores for the three *O*s have, therefore, been combined. In Fig. 2, the results are presented separately for each wavelength, the total point-scores for each color-name reported at that wavelength being given as percentages of the total possible score and organized into 3 × 3 tables.

Fig. 3 shows the scores for the longest-brightest and shortest-dimmest conditions at the various wavelengths. As expected from studies of the spectral location of the unique hues,[14] the maximal scores were found at 480 mμ for the blue, 520 mμ for the green, 576 mμ for the yellow, and at the longest wavelength tested for the red. The maximal white response was at 567 mμ (Fig. 2). In general, there was considerable stability of

λ	ft-L	Blue				Green				Yellow				Red				White			
	msec	.15	1.9	15	av	.15	1.9	15	av	.15	1.9	15	av	.15	1.9	15	av	.15	1.9	15	av
445	2	56	67	82	68									43	28	17	30	01	05	0	02
	20	52	68	84	68									48	30	15	31	0	02	01	01
	300	53	61	75	63									47	38	24	36	0	0	01	00
	av	54	66	81										46	32	19		00	02	01	
480	2	83	92	97	91	05	08	03	06									10	0	0	03
	20	96	88	92	92	02	12	07	07									02	0	0	01
	300	94	98	99	97	06	02	01	03									0	0	0	00
	av	91	93	96		04	07	04										04	00	00	
497	2	40	32	32	35	60	68	67	65												
	20	33	27	45	35	67	72	55	65												
	300	41	34	44	40	59	66	56	60												
	av	38	31	41		62	69	59													
520	2	14	07	01	07	81	87	93	87	0	06	06	04					05	0	0	02
	20	12	08	03	08	82	88	94	88	02	03	02	02					05	0	0	02
	300	12	0	0	04	87	92	93	91	01	07	07	05					0	0	0	00
	av	12	05	01		83	89	94		01	06	05						03	00	00	
546	2					58	72	72	67	17	21	26	21					25	07	02	11
	20					63	68	76	69	21	28	23	24					16	03	01	07
	300					67	69	74	69	25	34	26	28					07	01	0	02
	av					63	69	74		21	28	25						16	04	01	
567	2					02	15	12	10	50	79	86	72					47	06	02	18
	20					07	09	28	15	44	74	70	63					48	17	02	22
	300					13	27	36	26	53	72	64	63					33	0	0	11
	av					08	17	26		49	75	73						43	07	01	
576	2					0	0	02	01	80	97	97	92	02	01	0	01	18	02	0	07
	20					0	0	10	03	82	99	87	89	0	0	02	01	18	01	01	07
	300					0	0	09	03	77	98	91	89	0	02	0	01	22	0	0	08
	av					00	00	07		80	98	92		01	01	01		20	01	00	
585	2									62	64	72	66	34	36	28	33	03	0	0	01
	20									53	66	76	65	34	34	24	31	12	0	0	04
	300									60	67	77	68	37	33	22	31	03	0	0	01
	av									59	66	75		35	34	25		06	00	00	
598	2									47	53	62	54	47	47	37	44	06	0	0	02
	20									40	52	65	52	52	47	35	44	08	01	0	03
	300									46	53	67	55	51	47	33	44	03	0	0	01
	av									44	53	65		50	47	35		06	00	00	
605	2									25	25	30	27	72	72	70	71	03	02	0	02
	20									20	19	26	22	67	79	74	73	13	02	0	05
	300									27	31	33	30	72	69	67	69	02	0	0	01
	av									24	25	30		70	74	70		06	01	00	
655	2	12	12	14	13									88	87	86	87				
	20	12	12	12	12									88	88	87	88				
	300	12	14	09	12									88	86	91	88				
	av	12	13	12										88	87	88					

FIG. 2. THE PERCENTAGE-RESPONSE FOR EVERY COLOR REPORTED AT EACH WAVELENGTH AVERAGED FOR THE THREE NORMAL *O*S

color-naming through the range of conditions; the shapes of the functions for the various colors are quite similar under all conditions, and the peaks are invariant, although not much importance should be attached to the latter result, because the wavelengths tested are rather widely spaced.

---

[14] Dimmick and Hubbard, *op. cit.*, 242–254.

Certain changes are, nevertheless, apparent. First, except for the red, the curves usually are higher for the long-bright than for the short-dim condition. Secondly, there is a shift for both the red- and green-functions towards the shorter wavelengths under the short-dim condition; that is, there is an increase in green-responses (in the blue-green region) of the spectrum and in red-responses (in the yellow region) with a decrease in luminance and exposure-time. Variations produced by changes in luminance were much greater than those produced by variations in duration; there were only a few instances of sizable change as a function of exposure-time.

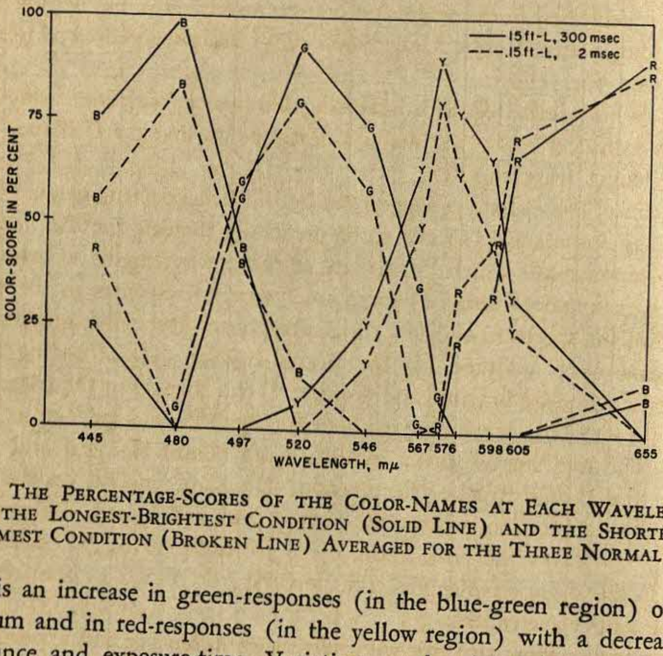


FIG. 3. THE PERCENTAGE-SCORES OF THE COLOR-NAMES AT EACH WAVELENGTH FOR THE LONGEST-BRIGHTEST CONDITION (SOLID LINE) AND THE SHORTEST-DIMMEST CONDITION (BROKEN LINE) AVERAGED FOR THE THREE NORMAL *O*s

The results may be summarized as follows: (1) The major change in the blue-score was a sharp increase as a function of increasing luminance at 445 mμ and a much smaller increase at 480 mμ for all *O*s. There was no linear change at 497 mμ, and it decreased at 520 mμ, suggesting a gradual shift in its behavior with increasing wavelength. As a function of increasing duration, blue was quite variable.

(2) Green increased both with increasing luminance and, to some ex-

tent, with duration unless it was 'mixed' with an appreciable amount of blue; then inverted *U*-shaped functions appeared. Under short-dim conditions, green-responses gave way to blue- and white-responses.

(3) Yellow, when 'mixed' with green, described an inverted *U*-shaped function with luminance, because at the low luminance white-responses increased, while at the high luminance green-responses increased. Without green, yellow increased with increasing luminance. As a function of duration, yellow was highly variable.

(4) Red was stable as a function of duration. With increasing luminance, it decreased in the violet and orange wavelengths, but became stable at the longer wavelengths. Under short-dim conditions, red-responses were substituted for yellow.

(5) White decreased with increasing luminance; with increasing duration, it decreased below 567 mμ and described an inverted *U*-shaped curve at the longer wavelengths.

Certain of these changes suggest the Bezold-Brücke phenomenon.[15] With increasing luminance (1) the scores generally increase for the dominant color—an increase which is smallest at those wavelengths at which the score is maximal[16]—and (2) there are fewer red-responses in the orange region of the spectrum and fewer green-responses in the blue-green region. The data were analyzed for the Bezold-Brücke shift according to the method devised by Boynton and Gordon.[17] The changes in the color-scores from 0.15-15 ft-L. were analyzed both for the 300- and the 2-m.sec. durations. The two curves shown in Fig. 4 are almost identical and clearly depict the Bezold-Brücke phenomenon. The same analysis on the color-changes as a function of exposure-time at 0.15 and 15 ft-L.—also plotted in Fig. 4—shows the peak in the green for both curves, but neither one rises in the short wavelengths and the curve for 0.15 ft-L. does not drop in the long wavelengths.

The results show also that with decreasing luminance and exposure-time the responses become more similar to those expected of tritanopes. Middleton and Holmes list the following as characteristics of tritanopia: confusions of yellow with gray, blue-green with blue, and orange with reddishpurple; yellow turns white at 562-570 mμ.[18] In this experiment, there were increased reports of blue at 520 mμ, red-orange tended to become more red, and there were increased reports of white in the yellow-greens and

---

[15] See D. McL. Purdy, Spectral hue as a function of intensity, this JOURNAL, 43, 1931, 541–559.

[16] There were no completely stable scores for yellow, green, or blue, although wavelengths quite close to the invariant hues were tested. It is interesting to note that the invariance of these hues does not hold for precise measurement (R. A. Weale, in *Colour Vision: CIBA Symposium*, A. V. S. de Reuck and Julie Knight (eds.), 1965, 365.

[17] Boynton and Gordon, *op. cit.,* 78–86.

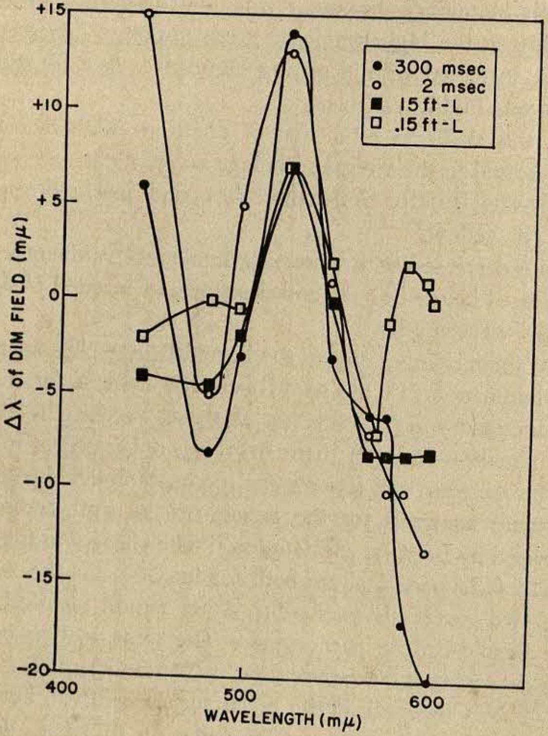


FIG. 4. THE BEZOLD-BRÜCKE HUE-SHIFT AS LUMINANCE IS INCREASED FROM 0.15–15 FT-L. FOR DURATIONS OF 300 (FILLED CIRCLES) AND 2 M.SEC. (OPEN CIRCLES) AND AS DURATION IS INCREASED FROM 2–300 M.SEC. FOR LUMINANCES OF 0.15 (OPEN SQUARES) AND 15 FT.-L. (CLOSED SQUARES)

especially at 567 mμ, precisely where specified. These effects were much more apparent with changes in luminance than exposure-time.

The results for the deuteranopic *O* are presented in Fig. 5, and the longest-brightest and shortest-dimmest conditions are plotted in Fig. 6. As expected, reports of yellow predominate throughout the spectrum above 520 mμ. It is only in the blue-green portion that the results resemble those for the normals, although the maximal score for blue occurred at 445 mμ rather than 480 mμ, and no red was reported at 445 mμ. The other maxima are at 520 mμ for green, 576 mμ for yellow, and 655 mμ for red. The red peak is only slightly greater than the score at 546 mμ and is clearly fortuitous, but the yellow peak at 576 mμ is quite sharp.

---

[18] W. E. K. Middleton and M. C. Holmes, The apparent colors of surfaces of small subtense—a preliminary report, *J. opt. Soc. Amer.*, 39, 1949, 582–592.

λ	ft-L / msec	Blue				Green				Yellow				Red				White			
		.15	1.9	15	av	.15	1.9	15	av	.15	1.9	15	av	.15	1.9	15	av	.15	1.9	15	av
445	2	100	93	100	98	0	0	0	00									0	07	0	02
	20	100	100	93	98	0	0	0	00									0	0	07	02
	300	100	100	97	99	0	0	03	01									0	0	0	00
	av	00	98	97		00	00	01										00	02	02	
480	2	67	80	87	78	20	20	13	18					07	0	0	02	07	0	0	02
	20	63	73	77	71	33	27	20	27					0	0	0	00	03	0	03	02
	300	77	70	90	79	23	30	10	21					0	0	0	00	0	0	0	00
	av	69	74	84		26	26	14						02	00	00		03	00	01	
497	2	50	50	53	51	27	50	47	44					07	0	0	02	07	0	0	02
	20	40	17	57	38	60	83	43	62					0	0	0	00	0	0	0	00
	300	30	40	67	46	57	60	33	50					13	0	0	04	0	0	0	00
	av	40	35	59		51	64	41						07	00	00		02	00	00	
520	2					77	70	47	64	13	23	30	22	10	07	23	13	0	0	0	00
	20					87	73	57	72	10	27	30	22	0	0	13	04	03	0	0	01
	300					70	63	77	70	20	37	23	27	07	0	0	02	03	0	0	01
	av					78	69	60		14	29	28		06	02	12		02	00	00	
546	2					53	33	03	30	37	60	60	52	10	07	37	18	0	0	0	00
	20					60	50	20	43	33	47	60	47	0	03	20	08	07	0	0	02
	300					37	33	13	28	47	57	53	52	17	10	33	20	0	0	0	00
	av					50	39	12		39	54	58		09	07	31		02	00	00	
567	2					30	27	0	19	50	60	53	54	10	07	47	21	10	07	0	06
	20					40	13	07	20	57	60	63	60	0	27	30	19	03	0	0	01
	300					20	10	07	12	47	67	70	61	33	23	23	27	0	0	0	00
	av					30	17	04		51	62	62		14	19	33		04	02	00	
576	2					07	0	03	03	83	90	90	88	10	03	07	07	0	07	0	02
	20					0	0	0	00	93	97	97	96	07	03	03	04	0	0	0	00
	300					0	0	0	00	80	90	90	87	20	10	10	13	0	0	0	00
	av					02	00	01		86	92	92		12	06	07		00	02	00	
585	2					27	20	10	19	60	67	70	66	13	13	20	16				
	20					17	20	0	12	77	70	83	77	07	10	17	11				
	300					20	20	0	13	67	67	73	69	13	13	27	18				
	av					21	20	03		68	68	75		11	12	21					
598	2					10	07	03	07	73	83	70	76	07	07	27	13	10	03	0	04
	20					27	03	03	11	57	90	80	76	07	07	17	10	10	0	0	03
	300					10	10	03	08	63	73	67	68	27	17	30	24	0	0	0	00
	av					16	07	03		64	82	72		13	10	24		07	01	00	
605	2					07	20	20	16	73	67	63	68	20	13	17	17	0	0	0	00
	20					17	27	0	14	67	60	83	70	13	13	17	14	03	0	0	01
	300					20	10	07	12	60	63	67	63	20	27	27	24	0	0	0	00
	av					14	19	09		67	63	71		18	18	20		01	00	00	
655	2					13	13	07	11	73	70	80	74	07	07	13	09	07	10	0	06
	20					17	03	03	08	70	80	80	77	07	17	17	13	07	0	0	02
	300					20	07	0	09	70	63	63	66	10	30	37	26	0	0	0	00
	av					17	08	03		71	71	74		08	18	22		04	03	00	

FIG. 5. THE PERCENTAGE-RESPONSE FOR EVERY COLOR REPORTED AT EACH WAVELENGTH BY THE DEUTERANOPIC *O*

Comparing these results with the changes occurring for normals at the same wavelengths, one can see that changes in luminance are somewhat more effective than changes in duration. Other than that, however, the

BUREAU OF EDNL. & PSYL. RESEARCH LIBRARY
J—407
8.9.70

deuteranope's responses are markedly different from those of the normals. For the deuteranope, green-responses decrease greatly with increasing luminance, but for the normals the reverse is true, and green is an inverted *U*-shape function of duration for the deuteranope, which never appears for normals. By contrast, yellow, which generally gives an inverted *U*-shape as a function of luminance for normals, simply increases for the deuteranope. Finally, red generally increases with increasing luminance for the deuteranope but decreases for normals when it is not stable.

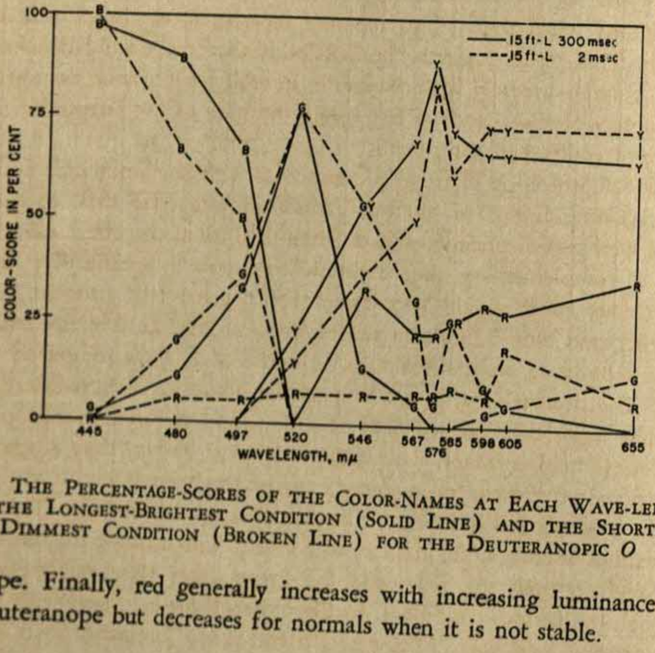


Fig. 6. The Percentage-Scores of the Color-Names at Each Wave-length for the Longest-Brightest Condition (Solid Line) and the Shortest-Dimmest Condition (Broken Line) for the Deuteranopic *O*

Discussion. Despite considerable stability of color-naming over the range of conditions tested, certain marked changes occur as the intensity and duration of the test-flash are varied. Some of these changes conform to changes previously found in measuring two-color thresholds with the method of constant stimuli; for example, in yellow-blue mixtures, the threshold for blue is minimal at the shortest exposure time.[19] In the present study, the number of reports of blue appeared to increase with decreasing stimulus-duration at 445 mμ and at the highest luminance of 655 mμ, the two wavelengths at which red was also perceived. Another previous finding is that, at short durations, green stimuli appear blue and yellow ones appear red. In the present results, it can be seen that from 585-605 mμ the red-scores tend to increase with decreasing luminance and duration, while at 520 mμ the blue-scores increase as luminance and duration decrease. The latter phenomenon did not, however, appear at the other blue-green wavelengths.

---

[19] Luria and Weissman, *op. cit.*, 1068–1072.

It is of considerable interest that the changes in color-naming exhibit at least part of the Bezold-Brücke effect as a function of changes in exposure-time when there is no apparent increase in the brightness of the stimuli. This result suggests the possibility that the Bezold-Brücke shift is not associated merely with changes in intensity, but may be demonstrable in a variety of ways, such as by varying the state of *O*'s adaptation or the degree of contrast of the stimuli.[20]

The demonstration of tritanopic responses with low luminance and short exposure-time also is of interest. Farnsworth suggested that, in addition to the well-known phenomenon of small-field tritanopia, there also should exist a complementary short-duration tritanopia.[21] Specifically, the tritanomalous effect should be constant for a given total product of area, intensity, and time. Middleton and Holmes did, in fact, find the tritanopic effect to be more pronounced at an exposure of 1/25 as compared with 5 sec.[22] The present data do not support the notion of a short-duration tritanopia, since the color-changes as a function of duration *per se,* with brightness held constant, were rather minimal. Rather, they suggest that any tritanopic effect is a consequence of the total energy involved.

The most marked individual differences occurred at 497 mμ. There were substantial reports both of blue and green, with neither predominating. For one *O,* both the blue- and the green-scores were relatively stable through the various conditions; for the others, blue increased with increasing duration and was *U*-shaped with increasing luminance, while green was the opposite.[23]

Table I gives an indication of the individual differences in this region of the spectrum. It shows the wavelengths at which the functions for the various color-names intersect for each *O* under the longest-brightest and shortest-dimmest conditions. These points are much more constant for the intersection of red and yellow, and green and yellow, than of blue and green. It seemed reasonable to assume that these differences were related to the relatively large variability of the wavelength at which 'unique' green has been determined by different investigators[24] and, indeed, is reported to be perceived by different *O*s. Rubin has found that not only is the *SD* of the green-point much larger than that of blue or yellow-points for color-normals, but that the frequency-distribution of the green points is bimodal.[25] Since the data suggested that the present *O*s might not have the same unique green-point, these were measured using stimulus-conditions similar to one of the experimental conditions: 50-min. test-stimulus, 200-m.sec. duration, 0.12 ft-L., and a white surround of equal brightness. No differences were found; all three *O*s reported unique green at about 507 mμ. The reasons for the differences in the color-naming results must, therefore, lie elsewhere. They may be due, for example, to marked individual differences in the temporal characteristics of the color-mechanisms, such as have recently been found by Kinney.[26]

---

[20] See also L. M. Hurvich and Dorothea Jameson, Some quantitative aspects of an opponent-colors theory: II. Brightness, saturation and hue in normal and dichromatic vision, *J. opt. Soc. Amer.* 45, 1955, 609–610.

[21] Dean Farnsworth, Tritanomalous vision as a threshold function, *Die Farbe,* 4, 1955, 185–196.

[22] Middleton and Holmes, *op. cit.,* 582–592.

[23] An inverted *U*-shape for green as a function of increasing luminance was found by Walraven, *op. cit.,* 50. He found it also for other colors using very small stimuli near threshold.

TABLE I

THE INTERSECTIONS OF THE COLOR-NAME FUNCTIONS FOR EACH *O* AT THE LONGEST BRIGHTEST (*LB*) AND SHORTEST-DIMMEST (*SD*) CONDITIONS

	Blue-green		Yellow-green		Red-yellow	
O	LB	SD	LB	SD	LB	SD
AR	492	489	555	550	596	594
DW	497	490	561	558	596	586
JK	502	522	557	560	596	588

The marked variations with duration in wavelengths which are found for yellow are not, however, due to individual differences. They appear to depend on whether or not yellow is the dominant color. For example, at 567 mμ, where yellow is dominant, the *O*s substitute white for green at low luminances; at 546 mμ, where green is dominant, they more often substitute white for the yellow.

Considerable attention has been given to the question of reciprocity between luminance and exposure-time.[27] The product of the two remains constant at threshold below a certain critical duration. This value is about 0.1 sec., although it varies as a function of wavelength for foveal stimuli of the size used in this study; blue stimuli continue to show summation much longer than do red.[28] Although these studies have been done at absolute threshold, Kinney and Connors have found that $It=C$ for measures of foveal acuity far above absolute threshold.[29] The question then may be asked if the present data give analogous evidence of constancy of color-naming when luminance × time is constant. Such conditions were, of course, set up empirically, since at each luminance-level of the surround the luminance of each stimulus of each duration was adjusted to match the surround in brightness. Thus, at each luminance-level there are three conditions (*i.e.* exposure-times) at which the visual effect is constant. Nevertheless, color-naming, while less variable with duration at a given luminance-level than between luminance-levels, was not, however, constant.

---

[24] Dimmick and Hubbard, *op. cit.*, 245.

[25] M. L. Rubin, Spectral hue loci of normal and anomalous trichromats, *Amer. J. Ophthal.*, 52, 1962, 166–172.

[26] J. A. S. Kinney, Color induction using asynchronous flashes, *Vision Res.*, in press.

[27] N. R. Bartlett, Thresholds as dependent on some energy relations and characteristics of the subject, in *Vision and Visual Perception*, C. H. Graham (ed.) 1965, 170–174.

There has been little work on color-naming by color-defectives. What has been done appears to center on the problem of differential chromatic contrast between color-normals and defectives.[30] Attention has been paid to color-naming by unilateral color-defectives, but the investigators generally have been content simply to survey the spectrum.[31]

The present results for a deuteranopic *O* show few similarities to the results for the color-normals. It is agreed that deuteranopes essentially see only blue, yellow, and a colorless neutral point. On this basis, the green-function shown by the deuteranope must be the result of confusion between green and gray. In an effort to clarify this point, the deuteranope was asked to make further observations with the stimuli from 480-546 mμ as well as a 'white' stimulus at the longest-brightest and shortest-dimmest conditions.[32] The white stimulus was called 'green' 63 and 67% of the time, respectively, indicating confusion of green and white. The 'green'-function for the chromatic stimuli was virtually unchanged from the original data for the longest-brightest condition but was somewhat lower and less sharply peaked for the shortest-dimmest condition; there also were more responses of 'yellow.' Since, however, the white stimulus evoked more responses of "white" than did any of the chromatic stimuli, the deuteranope must have been able to perceive some difference between the white and green stimuli, possibly the yellowish cost of the white stimulus.

---

[28] H. G. Sperling and C. L. Jolliffe, Intensity-time relationship at threshold for spectral stimuli in human vision, *J. opt. Soc. Amer.* 55, 1965, 191–199.

[29] J. A. S. Kinney and M. M. Connors, Recovery of foveal acuity following exposure to various intensities and durations of light, this JOURNAL, 78, 1965, 432–440.

[30] Chapanis, Simultaneous chromatic contrast in normal and abnormal color vision, this JOURNAL, 62, 1949, 526–539.

[31] C. H. Graham, H. G. Sperling, Yun Hsia, and A. H. Coulson, The determination of some visual functions of a unilaterally color-blind subject: Methods and results, J. Psychol. 51, 1961, 2–32, 489; D. B. Judd, Basic correlates of the visual stimulus, in *Handbook of Experimental Psychology,* in S. S. Stevens (ed.), 1951, 827–836.

[32] The tungsten lamp plus a Wratten 78 filter produced a slightly yellowish white.

The sharp peak in the yellow-response at 576 mμ, coinciding with the peak for the color-normals, was somewhat unexpected. Sloan and Wollach, in their study of a unilateral deuteranope, did, however, report that the only wavelengths which were identical in appearance to both eyes of their *O* were those chosen as unitary yellow and unitary blue,[33] while Graham and co-workers reported different degrees of saturation for yellow and blue at different wavelengths.[34] Thus, although it it accepted that the deuteranope sees only yellow above the neutral point, there apparently are different qualities of yellow, and the 'most' yellow wavelength, even for deuteranopes, is that which normals choose as unique yellow. Similarly, the deuteranope's peak blue-response in the present study did not coincide with the normal peak (480 mμ), but conformed more closely to the wavelength chosen by Sloan and Wollach's *O* as unique blue—452 mμ[35]—and to the findings by Rubin that the saturation of blue increases with decreasing wavelength for deuteranopes.[36]

Summary

By the method devised by Boynton, Shafer, and Neun, three color-normals and one deuteranope named the colors of 11 wavelengths from 445-655 mμ which were varied from 0.15-15 ft-L. in intensity and from 2-300 m.sec. in exposure-time. Despite considerable stability of color-naming over this range of conditions, certain marked changes occurred, more as a result of changes in luminance than in duration. These changes exhibited the Bezold-Brücke phenomenon, not only as a function of intensity but to some extent as a function of duration with brightness held constant. The results also showed that with decreasing luminance and exposure time, the *O*'s responses became similar to a 'low-energy tritanopia' as well as the well-known 'small-field tritanopia.'

The maximal score for the four colors—blue, green, yellow, and red—

---

[33] L. L. Sloan and Lorraine Wollach, A case of unilateral deuteranopia, *J. opt. Soc. Amer.* 38, 1948, 502–509.

[34] Graham, Sperling, Hsia, and Coulson, *op. cit.*, 24.

[35] Sloan and Wollach assumed that this choice of unique blue resulted from abnormal ocular pigmentation. If that were true also of the present deuteranope, then presumably he should have shown a maximal yellow-response above 572 mμ.

[36] Rubin, *op. cit.*, 169.

conformed to determinations of the spectral locations for the 'unique' colors. At these points, the scores always were higher for the brightest-longest condition than for the shortest-dimmest condition. The main change in the blue-scores was a sharp increase with increasing luminance and duration, while red generally decreased. Yellow showed certain reliable variations between wavelengths. The only marked individual differences occurred near 500 mμ, but they did not correlate with measurements of the unique green-point. The results for the deuteranopic *O* were markedly different from those of the normal *O*s in nearly all respects.

THE SUBTHRESHOLD PERCEPTION OF STIMULUS-MEANING

By JOHN A. MCNULTY, FREDERICK J. DOCKRILL, and BETTY ANN LEVY,
Dalhousie University

This paper is concerned with the possibility that subjects (*Ss*) may be able to respond as if they have perceived the meaning of tachistoscopically presented words before they are aware of the physical characteristics (*i.e.* the actual letters) of the words. Although the bulk of experimental evidence supports the notion that the physical structure of the stimulus-word is the most important source of information for tachistoscopic identification, there are findings which suggest that meaning also affects thresholds. In both cases, some kind of a partial identification-process is usually postulated. Partial identification refers to the perception of partial cues from stimuli presented at very brief exposure-durations before complete identification-levels have been reached. The first suggestions that such a process existed arose from the early studies in perceptual defense and subception.

For example, Bricker and Chapanis found that, after an incorrect verbal response to a tachistoscopically exposed stimulus, the mean number of additional guesses required by *S* to name the stimulus correctly was significantly less than would be expected on the basis of random guessing.[1] They postulated that fragmentary cues, perceived during the presentation of the stimulus, limited the number of plausible responses. The cues were thought to be syllables and letters which stood out in the brief stimulus-exposure, thus limiting the responses that resembled the cues perceived.

Numerous other studies have been reported which support similar conclusions.[2] Eriksen, Azuma, and Hicks found that, even when *Ss* could not identify the stimulus-words, they could often make affective judgments regarding their pleasantness or unpleasantness,[3] which were correct significantly more often than would be expected by chance. There was no evidence, however, that these judgments lowered the thresholds for complete identification of the items. Thresholds were no lower than those of items for which *S*s could make no affective judgments.

---

* Received for publication July 7, 1966. This research was supported by Grants APT- and APB-86 from the National Research Council of Canada.

[1] P. D. Bricker and Alphonse Chapanis, Do incorrectly perceived tachistoscopic stimuli convey some information?, *Psychol. Rev.*, 60, 1953, 181–188.

[2] Cf., for example, the following studies; A. E. Edwards, Subliminal tachistoscopic perception as a function of threshold method. *J. Psychol.*, 50, 1960, 139–144; F. R. Goldberg and H. Fiss, Partial cues and the phenomenon of "discrimination without awareness," *Percept. mot. Skills*, 9, 1959, 243–251; B. B. Murdock, Jr., Perceptual defense and threshold measurements, *J. Pers.*, 22, 1954, 565–571; Jan Pierce, Some sources of artifact in studies of the tachistoscopic perception of words, *J. exp. Psychol.*, 66, 1963, 363–370; Morton Weiner and P. H. Schiller, Subliminal perception or perception of partial cues, *J. abn. soc. Psychol.*, 61, 1960, 124–137.

[3] C. W. Eriksen, Hiroshi Azuma, R. B. Hicks, Verbal discrimination of pleasant and unpleasant stimuli prior to specific identification. *J. abn. soc. Psychol.*, 59, 1959, 114–119.

In a perceptual defense study, Dixon found that verbal responses to subliminal stimuli often tended to be related autistically or symbolically to the stimulus-material.[4] He interpreted this finding as indicating that meaning was the primary determinant of response. Fuhrer and Eriksen failed to replicate Dixon's results and stressed the importance of structural cues.[5] Recently, Worthington has argued that Fuhrer and Eriksen's results may be due to procedural artifacts resulting from their failure to replicate exactly crucial features of Dixon's procedure.[6] Worthington measured 'dark adaptation'—the time required for *S*s to indicate the presence of a faint white light (actually the stimulus-word) after a pretest-period of light-adaptation. His results indicated that the rate of dark adaptation to "taboo" words was significantly longer than to neutral words, to the 'taboo' words inverted, or to words structurally similar to the emotional stimuli. Worthington's experiment suggests, therefore, that the *S*s were responding to the meaning of the stimuli before they could report anything about the structure of the stimuli, because they showed perceptual defense even when they were unaware that words were being presented. His design also eliminated any possibility that these results could be interpreted in terms of response-suppression, since the *S*s did not have to say any of the taboo words but only had to indicate their presence.

Our experiments were designed to test further the implication that our *S*s may be able to respond in a way which indicates they have perceived the meaning of a stimulus before they are aware of its structure. In the first two experiments, a method similar to that used by Ross, Yarczower, and Williams was employed.[7] On a pretrial, their *S*s learned 10 stimulus-words presented at suprathreshold-exposures. On a later test, words varying in their degree of relationship to the pretrial-words were presented on the tachistoscope and identification-thresholds were obtained. The assumption was made that the words learned on the pretrial would tend to be elicited when words similar to them were presented on the tachistoscope, causing response-interference and therefore higher thresholds for such words. Althought the results of the study are not directly applicable to the present problem, since meaning was not one of the variables investigated, the method is. By use of this method, words of high associative similarity, *i.e.* of similar meaning, to the pretrial-words can be presented tachistoscopically. If these words are so selected that their letter structure is different from that of the pretrial-words, and, if structural cues have to be perceived before the *S*s can make judgments about the meaning of the stimuli, then such test-words should have identification-thresholds no different from words that have no relation whatever to the pretrial-words. This follows from the fact that as soon as *S* perceives the word's structure (perhaps even a part of it) he will know that such a test word is not a word from the pretrial list. If, on the other hand, meaning can be perceived before *S* can make a response indicating that he has perceived the structure of the word, then the pretrial words would tend to be elicited when words of high associative similarity to them are presented tachistoscopically. As a result, these test-words should have higher thresholds than words unrelated to the pretrial-words.

---

[4] N. F. Dixon, The effect of subliminal stimulation on automatic and verbal behavior, *ibid.,* 57, 1958, 29–36.

[5] M. J. Fuhrer and C. W. Eriksen, The unconscious perception of the meaning of verbal stimuli, *ibid.,* 61, 1960, 432–439.

[6] A. G. Worthington, Differential rates of dark adaptation to "taboo" and "neutral" stimuli, *Canad. J. Psychol.,* 18, 1964, 257–265.

[7] Sherman Ross, Matthew Yarczower, and G. M. Williams, Recognitive thresholds for words as a function of set and similarity, this JOURNAL, 69, 1956, 82–86.

TABLE I

STIMULUS-WORDS USED IN THE EXPERIMENT

List A-1	T-L count*	Association-value between A-1 and A-2 Words†	List A-2	T-L count*	List B	T-L count*	Practice-words	T-L count*
drowsy	10	2.55	sleepy	15	nimble	6	uphill	3
showy	2	2.49	gaudy	5	balmy	5	mundane	1
select	A	2.48	chosen	A	hollow	A	fervent	4
limber	1	2.17	supple	3	candid	3	erect	44
sallow	3	1.86	pallid	3	gurgle	3		
yellow	AA	1.83	golden	AA	modern	AA		
shaky	3	1.82	jerky	1	murky	1		
brittle	6	1.58	fragile	7	wistful	7		
intact	4	1.57	unhurt	2	hectic	2		
mystic	4	1.47	occult	2	porous	2		

* Thorndike-Lorge word-count.

† From Melton and Safier, unpublished study cited in Hilgard, *op. cit.*, 547–552.

EXPERIMENT I

Method. Four practice-words and three lists of 10 stimulus-words each were used in the experiment. All of these words are shown in Table I. Ten pairs of two-syllable adjectives with established associative values between the two pair members were taken from an unpublished study by Melton and Safier.[8] The pairs were chosen on the basis of the following criteria: (1) that each member of a particular pair be of the same number of letters in length; (2) that each pair be of approximately the same relative frequency of occurrence in printed English according to the Thorndike and Lorge general word-count;[9] (3) that the two pair members be of different over-all letter structure; and (4) that only those pairs be selected that had a high associative value between the two pair members. One member of each pair thus selected was randomly assigned to List A-1 and the other to List A-2.

---

[8] This study is described in E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 547–552.

List B in Table I was selected from the Melton and Safier lists to match Lists A-1 and A-2 on criteria (1) to (3) above. In addition, there had to be no obvious association between the matched words of List B and those of Lists A-1 and A-2.

In none of the lists was it possible to control completely for duplication of intra-item letter sequences due to the limited number of word pairs that met the required criteria. This meant that there was partial duplication of structure from List A-1 to A-2 in some words (*e.g.* the final "y" in "showy" from List A-1 also appeared in "gaudy" from List A-2). It should be noted, however, that there was just as much duplication in the control list, List B (*e.g.* the word from List B, "blamy," chosen to match "showy" and "gaudy" also had a final "y").

The practice words were randomly selected from the Thorndike-Lorge word-count with the restrictions: (1) that the words be two-syllable adjectives not used in Lists A-1, A-2, or B; (2) that there be no obvious association between them and the words in Lists A-1, A-2, and B; and (3) that the words vary in length from five to seven letters (the same range as in the other lists).

Each of the stimulus-words was typed in the center of a white, unlined, 3 × 5 in. index-card in lower case letters. The words were typed with an IBM electric typewriter equipped with a carbon ribbon to insure that the intensity and shading of letters would be uniform within and between words. The cards were presented on a Gerbrands Harvard Tachistoscope adapted to hold them so that the words appeared in the center of the stimulus-screen.

Subjects. Thirty *S*s, 15 men and 15 women, enrolled in an undergraduate course in experimental psychology took part in the experiment. They were randomly assigned to three groups of 10 *S*s each, with the restriction that each group contain 5 men and 5 women.

Procedure. The experiment was conducted in a darkened room, the only light being provided by a 40-w. goose-neck lamp illuminating the control unit of the tachistoscope. The words were all presented in the tachistoscope at a constant level of illumination that was used throughout the entire study—in Experiments II and III, which follow, as well as in Experiment I.

Using one of the four practice-words, *S* was first acquainted with the manner of operation of the tachistoscope. Then he was presented with one of the lists of 10 words to learn. Items in the list were presented on the apparatus at the rate of 2 sec. per item. After the list had been presented, *S* was allowed 1 min. during which he attempted to recall all of the words in the list. Following this, the items were presented for a second trial. This procedure was continued until *S* had learned the list to a criterion of two consecutive perfect recalls. A different random order of the 10 words was presented on each trial.

The list that the *S*s were given to learn depended upon the group to which they had been assigned. Group I learned List A-1; Group II learned List B; and Group III learned List A-2.

---

[9] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

Once the learning criterion had been reached, *S* was allowed to relax for 3 min. and the second part of the experiment was begun. The other three practice-words (the fourth had already been used to acquaint *S* with the apparatus) were then used to give *S* practice in identifying words presented at very short exposures. The three practice-words were presented repetitively in the same order at gradually increasing durations until the words could all be completely identified.

The test-list, A-2, was then presented. All the *Ss*, regardless of the group to which they had been assigned, received the same test-list. This meant that Group I was presented with a list containing words of high associative similarity to the words learned in the first part of the experiment. Group II received a list that was unrelated to the list presented in the learning session. Group III received a list identical to the one which they had previously learned.

For each *S*, *E* decided upon an initial exposure for the test-words that was just below that at which the practice-words had been identified. The entire test-list was presented at this exposure and then an ascending method of limits was used with exposures increasing in units of 0.01 sec. on successive presentations of the test-list. When a word was identified correctly the first time, it was exposed at the same duration on the next presentation of the list. The identification-threshold of any word was defined as that exposure at which the word was correctly identified twice in succession, although other words normally intervened between the two successive exposures of any particular word. This procedure was continued until identification-thresholds had been obtained for all of the words. When the identification-threshold had been obtained for any particular word, it was exposed at arbitrary durations on succeeding presentations. The order in which the individual test-words were presented was randomly varied from one presentation of the list to the next. After each complete presentation of the list, there was a 30-sec. break while the stimulus-cards were re-ordered for the next presentation.

TABLE II

MEAN THRESHOLDS OF IDENTIFICATION FOR THE THREE GROUPS

(In m. sec.)

	Group		
	I	II	III
$\overline{X}$	132.0	133.3	47.0
SD	94.18	52.52	48.16

Results. The means and *SD* of the thresholds of identification for the three groups of *Ss* are shown in Table II. As the table shows, thresholds for Group I, in which the test-list was of high associative similarity to the list learned previously, were about the same as those for Group II, in which the test list was unrelated to the learned list. Thresholds for Group III, in which the test-list was identical to the learned list, were significantly less than those for the other two groups ($t > 2.54$; 18 *df;* $p < 0.025$).

In Group III, partial cues and information received by the *Ss* during tachistoscopic presentation of the test-words would correspond exactly to words learned prior to the test-session. Hence, thresholds are lower because *S*s need less information in order to identify the stimuli correctly.

The fact that there were no differences between Groups I and II suggests that the *S*s have to be aware of stimulus-structure before they can make a response indicating that they have perceived stimulus-meaning. If this were not so, then any subthreshold-information the *S*s in Group I had obtained about stimulus-meaning would have tended to elicit the words of high associative similarity which had been learned earlier, causing response-interference and higher thresholds than in Group II. Subthreshold-information about meaning would not have caused such interference in Group II since the test-words were unrelated to the words learned in the first part of the experiment.

EXPERIMENT II

Because variability among the *S*s was high, as the standard deviations in Table II indicate, there is the possibility that 'real' differences between Groups I and II are obscured by idiosyncratic factors associated with the *S*s. This is unlikely since an analysis of variance performed on the identification-thresholds for the practice-words showed that there were no significant differences among the three groups. Still, because the possibility does exist, the experiment was replicated, but all three types of test-words were presented to the same *S*s. That is, each *S* in Experiment II was given some words in the tachistoscope which were of high associative similarity to words he had learned previously, some words which were unrelated to the learned list, and some words which were identical.

Method. The materials and procedure used in this experiment were identical to those employed in Experiment I, with the following exceptions. Ten *S*s, 5 men and 5 women, took part in the experiment. They were drawn from the same psychology class as the *S*s in Experiment I. Again, each *S* first learned one of the lists of 10 words to a criterion of two consecutive perfect recalls. Half of *S*s learned List A-1 and the other half learned List A-2.

Following the same procedure as before, a number of test-words were then presented tachistoscopically to each *S*, and thresholds of identification for each word were obtained. This time, however, 15 test-stimuli were presented on the tachistoscope. The 15 words were made up of 5 words randomly chosen from List A-1, the 5 words in A-2 that were similar in meaning to the other 5 words of List A-1 that were not chosen, and 5 words randomly chosen from List B. The only restriction to the random selection of words was that each word be used equally often by the 10 *S*s. The order in which the words appeared in the tachistoscope was randomly varied from presentation to presentation with the restriction that no more than two words from a given list appear consecutively in the tachistoscope. Thresholds were determined as before, the only difference being that the first five words for which identification-thresholds were obtained were removed from the list on subsequent presentations. Ten words were always presented. Preliminary work showed that presenting 15 words each time was too fatiguing for *S*s and made this part of the experiment much longer than the comparable part of Experiment I. Removing the first five words that were identified kept the time for the tachistoscopic task approximately constant for the two experiments.

Presenting five words from each list in the tachistoscope meant that each *S* was given five words of high associative similarity to words he had learned previously (Condition I), five words unrelated to the learned list (Condition II), and five words that were identical to words in the learned list (Condition III). For the *S*s who had learned List A-1, the Condition I words were the five A-2 words, the Condition II words were the five B words, and the Condition III words were the five A-1 words. For *S*s who had previously learned list A-2, the Condition I words were the five A-1 words, the Condition II words were the five B words, and the Condition III words were the five A-2 words.

TABLE III

MEAN THRESHOLDS OF IDENTIFICATION FOR THE THREE TYPES OF WORDS (In m. sec.)

	Condition		
	I	II	III
$\overline{X}$	111.6	118.8	53.8
SD	60.94	64.17	51.17

Results. The mean and *SD* of the thresholds of identification for the three conditions are shown in Table III. In agreement with the results of Experiment I, the thresholds for words of high similarity to a list learned previously (Condition I) were about the same as those of words unrelated to the learned list (Condition II), and a *t*-test for matched groups indicated that the difference between them was not, of course, significant. Thresholds for words identical to words in the learned list (Condition III) were significantly less than those for the other two groups ($t > 2.69$; 9 *df*; $p < 0.025$).

These results again support the conclusion that *S*s have to be aware of stimulus-structure before they can respond to stimulus-meaning. If they did not have to, then the thresholds for words similar in meaning to previously learned words should have been higher than for either of the other two types of words, since the perception of even partial information about the meaning of these stimuli would have tended to arouse the words of similar meaning learned previously.

EXPERIMENT III

There is still the possibility, however, that the basic assumption underlying both Experiments I and II does not hold. The assumption is that if stimulus-meaning is perceived before the *S*s are aware of stimulus-structure, then words similar in meaning to previously learned words would tend to elicit these learned words, causing interference and higher thresholds than for unrelated words. There is still the possibility, however unlikely, that the *S*s did perceive stimulus-meaning before they were aware of stimulus-structure, but that this perception of meaning did *not* tend to elicit words of similar meaning learned earlier. Therefore, Experiment III was designed to eliminate this possibility by using a procedure which did not require *S*s to learn a list of words before being presented with words in the tachistoscope.

Method. In Experiment III, the method of recognition was employed. Words were presented in the tachistoscope for an arbitrary number of durations. Before each word was presented, *S* was given a card upon which two words were written that were of equal length and word-frequency but of different letter structure. If the two words on the card are similar in meaning to each other, and if stimulus-meaning can be perceived before *S* is aware of stimulus-structure, then, at exposure-durations at which *S* can perceive the meaning but is unaware of the structure of the stimulus, he will be confused as to which of the words was the one presented in the tachistoscope. Consequently, he is more likely to make a mistake. This kind of confusion would not occur if the two words on the card are unrelated in meaning to each other. That is, if subthreshold-perception of meaning is possible, then the number of words correctly recognized at such exposure-durations will be less when the alternatives of the recognitive test are highly similar in meaning to each other than when they are not. This discrepancy should decrease as the duration of the exposure of the words is increased, since more structural cues will be perceived as the length of the exposure is increased, and as soon as *S* can perceive the structure of the word he will know which of the two words on the card was the one presented on the tachistoscope. If, on the other hand, *S* must be aware of stimulus-structure before any information about stimulus-meaning is perceived, then the number of words correctly recognized at each exposure-duration will be the same when incorrect alternatives in the recognitive test are similar in meaning to the correct items as when they are unrelated to the test items.

Additional words in each list were needed for this experiment. These words are shown in Table IV. The words in each list were selected according to the same criteria as words from the corresponding lists of Experiment I. The words in Table IV were combined with the corresponding lists of Table I, giving 20 words in the A-1 List, 20 in the A-2 List, 20 in List B, and 8 practice words.

Twenty *S*s, 10 men and 10 women, from the same subject-population as that in Experiments I and II were used in this experiment. None had previously participated in Experiments I or II.

Following the same procedure as before, one of the practice-words was randomly selected to acquaint *S* with the manner of operation of the tachistoscope. Unlike Experiments I and II, however, he was not then given a list of words to learn. He was immediately presented with three of the practice-words in the tachistoscope. Each word was presented once at each of four exposure-durations: 10, 20, 40, and 100 m.sec. Before each word was presented, *S* was given a card with two words typed on it. He was told that one of the words was the one which was going to be presented in the tachistoscope and after the word had been presented he was to say which one it was. The other word on the card was another practice-word of equal length and approximately equal frequency to the one being presented. The incorrect alternatives on the cards were not other words which were to be presented in the tachistoscope. For the recognitive task, the first practice-word in Table I was combined with the first practice-word in Table IV, the second in Table I with the second in Table IV, and so on.

TABLE IV

ADDITIONAL WORDS USED IN EXPERIMENT III

List A-1	T-L count*	Association-value between A-1 and A-2 Words†	List A-2	T-L count*	List B	T-L count*	Practice words	T-L count*
rocky	30	2.57	stony	12	dizzy	10	abrupt	6
gloomy	19	2.49	dismal	13	decent	18	slavish	1
active	A	2.33	lively	36	secret	A	austere	5
godly	1	2.28	pious	12	milky	5	alike	43
divine	41	2.24	sacred	38	savage	40		
misty	9	2.15	foggy	4	comic	7		
thorny	4	1.85	barbed	1	greasy	4		
exempt	6	1.78	immune	1	tropic	9		
pretty	AA	1.77	lovely	A	direct	AA		
sinful	4	1.45	guilty	38	faulty	4		

* Thorndike-Lorge word count.

† From Melton and Safier, unpublished study cited in Hilgard, *op. cit.*, 547-552.

After the practice session had been completed, each *S* was next presented with 20 words in the tachistoscope at each of the four exposure-durations: 10, 20, 40, and 100 m.sec. The words were presented one at a time and the entire list of 20 words was presented first at 10 m.sec., then the entire list was presented at 20, next at 40, and finally at 100 m.sec. The order of presentation of the words was randomly varied for each exposure-duration. There was a 2-min. rest-period between successive presentations of the 20 words.

The 20 words that any *S* received consisted of 10 words randomly selected from List A-1 plus the 10 words from List A-2 that were unrelated to the particular 10 words selected from the A-1 List. The only restriction to the random selection of words was that each of the A-1 and A-2 words be used equally often over the whole group of *Ss*.

After the word was presented in the tachistoscope, *S* had to report which of the two words he saw. If he did not know, he had to guess. This procedure was continued until all the words had been presented at each duration of exposure.

Two types of recognition were used. For a recognitive task in which correct and incorrect alternatives were unrelated to each other, 10 of the 20 words which had been selected for any *S* were taken randomly and paired with their corresponding B-items. Each of these pairs was typed on a separate card to be given to *S* when the appropriate word was to be presented in the tachistoscope. For a recognitive task in which correct and incorrect alternatives were of high associative similarity to one another, the other 10 items of the list were paired with their corresponding words in either the A-1 or A-2 Lists. Each of these pairs was similarly typed on a separate card and given to *S* at the appropriate time. This meant that, of the 20 words presented to any *S*, 10 of them had to be recognized from cards which contained the word presented in the tachistoscope, plus a word unrelated to it in meaning or structure but of similar length and frequency of occurrence; and 10 of them had to be recognized from cards which contained the word presented in the tachistoscope, plus a word similar in meaning to it and of similar length and frequency but of different structure.

Results. The percentage of words correctly recognized at each duration employed is shown in Fig. 1. The 'unrelated' curve refers to words that were recognized from cards containing the correct word plus an alternative unrelated to it. The 'high association' curve refers to words that were recognized from cards containing the test-word plus another word of high associative similarity to it. As Fig. 1 shows, there was practically no difference between the two groups of words at any duration of exposure.

These results again support the conclusion that stimulus-meaning cannot be perceived before *S*s are aware of the structure of the stimulus. If it could be, then there would have been fewer items recognized correctly when the correct and incorrect alternatives in the recognitive task were similar in meaning to one another than when they were not. This difference would have decreased as exposure-duration increased and the *S*s became aware of stimulus-structure. Since the correct and incorrect alternatives on the recognitive tasks were always different in structure from one another, as soon as *S* was aware of the structure (or part of it) of a word, he would have no difficulty in determining which of the two items on the recognitive task was the correct word.

One might argue that all of the exposure-durations used in this experiment were high enough to yield at least partial information about stimulus-structure, and that if shorter durations had been used, the *S*s might still have obtained information about stimulus-meaning without being aware of any structural features of the stimulus. This seems unlikely, considering that at the shortest duration used, 10 m.sec., the *S*s were performing at almost the chance level. (With only two alternatives on the recognitive test for each test-word presented, the *S*s would have been correct about 50% of the time even if they had never looked at the words being presented in the tachistoscope.) This indicates that *Ss* were receiving very little information of any kind at the 10-m.sec. exposure.

DISCUSSION AND CONCLUSIONS

None of the experiments reported here yielded any evidence that stimulus-meanings can be perceived at subthreshold-levels. The results of the experiments support the view that the *Ss* must be aware of the physical structure of the stimulus before they can perceive its meaning. As Eriksen has pointed out, the meaning of a stimulus is conveyed by its physical characteristics and unless the *Ss* are aware of these characteristics the perception of meaning is impossible.[10]

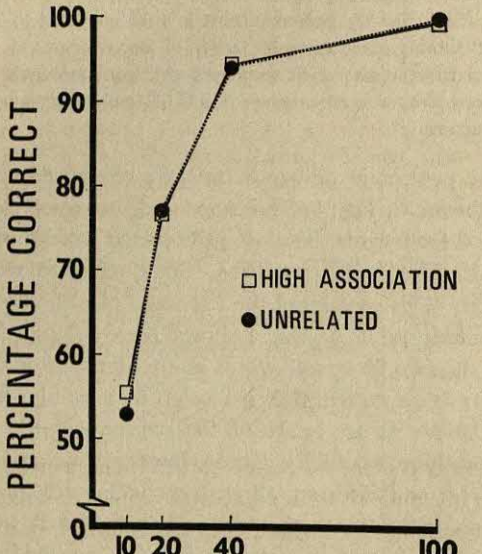


FIG. 1. PERCENTAGE OF WORDS CORRECTLY RECOGNIZED AT EACH DURATION OF THE EXPOSURE

The curve labelled "unrelated" refers to words recognized from cards containing the test-word plus another word unrelated to it. The curve labelled "high association" refers to words recognized from cards containing the test-word plus a word of high associative similarity to it.

On the basis of our study, we can only conclude that, in studies such as Worthington's which have purported to show the perception of meaning without awareness of stimulus-structure, the *S*s must in fact have been aware of the physical characteristics of the stimulus-words.[11] In all such studies, 'unawareness' of the stimuli was always inferred from *S*s' verbal reports. Since verbal reports are often unreliable, we tried so to design our experiments that it would not be necessary to rely upon verbal reports. Instead, we simply tried to construct situations in which certain specific results would be predicted if perception without awareness was possible and in which other predictable results would be obtained if it was not.

---

[10] C. W. Eriksen, Discrimination and learning without awareness: a methodological survey and evaluation, *Psychol. Rev.*, 67, 1960, 279–300.

All of our results support the position that the most important source (and perhaps the only source) of information for the perception of tachistoscopically presented words is their physical structure. Partial cues are probably perceived at brief exposures but, if so, they are undoubtedly structural cues. Using a recognitive procedure, Murdock showed, for example, that as the *S*s approached the recognitive threshold for a particular item, their responses were not distributed randomly among the various alternatives but rather were restricted more and more to alternatives which were structurally similar to the item being presented.[12] In terms of our experiments, one should note that partial cues of this type, unless they result in *S*'s 'guessing' the correct word, are unlikely to convey correct information about the meaning of that word. Consider, for example, the quite different meanings of words such as "bold" and "cold", or "interior" and "inferior", which are highly similar in structure to each other. It is our view, therefore, that meaning cannot be perceived until *S* is aware of the entire structure of the word and can identify or recognize it correctly, or until he is aware of enough of the word's structure so that he can guess it correctly. Only when this stage has been reached will the meaning of the stimulus be perceived.

SUMMARY

Three experiments were conducted to determine whether any information about the meaning of a stimulus could be perceived at exposures too brief to permit complete identification. In Experiment I, the *S*s were given a list of words to learn. Following this, a list of words similar in meaning, but of different structure, were presented tachistoscopically to one group of *S*s. A second group received words unrelated in meaning or structure to the words earlier learned. A third group received the same words as they had previously learned. It was predicted that, if meaning could be perceived at a subthreshold-level, then the words similar in meaning to words learned previously would tend to evoke the learned words, causing response interference. Therefore, *S*s given these words should have higher thresholds than those presented with words unrelated to the previously learned list. If subthreshold-perception of meaning does not occur, then the thresholds of the words similar in meaning to the learned words should be no higher than those of unrelated words. Results of the experiment

---

[11] Worthington, *op. cit.*, 257–265.
[12] Murdock, *op. cit.*, 565–571.

showed that there was no difference in identification-thresholds between the group of *S*s who were presented with words similar in meaning to the learned list and the group who received words unrelated in meaning to the words learned earlier. Thresholds for *S*s who were given words in the tachistoscope which were identical to the learned words were much lower than those of the other two groups.

The second experiment was a replication of the first with the exception that all three types of words were presented in the tachistoscope to each *S*. Results of this experiment were similar to those of the first. Identification-thresholds for words similar in meaning to previously learned words were about the same as those for unrelated words. Again, words identical to words learned earlier had the lowest tachistoscopic thresholds.

The third experiment employed a recognitive rather than an identification-procedure. A number of words were presented in the tachistoscope to each *S*. All of the words were presented once at each of four fixed durations. Before each word was presented, *S* was given a card which had on it the word to be presented in the tachistoscope plus another word. After each word had been presented, *S* had to indicate which of the two words on the card was the one he saw. The two words on the card were sometimes similar in meaning to each other and sometimes were unrelated. If subthreshold-perception of meaning is possible, fewer words should be correctly recognized when the alternatives on the cards are highly similar in meaning to one another than when they are not. The results of the experiment did not support this prediction. The number of words recognized at each exposure-duration was the same when the words had to be recognized from cards containing two words similar in meaning as when the cards contained two alternatives unrelated in meaning.

On the basis of these experiments it was concluded that *S*s must be aware of the physical structure of a stimulus before its meaning can be perceived.

THE PERCEPTION OF WHERE A PERSON IS LOOKING

By MARVIN G. CLINE, Howard University

The complex visual pattern that is the face of a person carries significant information about the person and the relationship he bears to his environment. On the one hand, this information alone tells the *O* something about the momentary state of the person.[1] On the other hand, the expression of the face helps to specify the emotion of the relationship, what Michotte has called the functional connection of the person to his environment.[2] If knowledge of the environment to which the person is related is also available to *O*, then the dispositional properties of the person begin to be revealed.[3] A particular pattern of the face, for example, may be the appropriate stimulus for the perception of anger in the person. To know the object in the environment which provoked that anger, and to observe how the angry person behaves toward the object, is to acquire information about the kind of person who is exhibiting the anger.

Clearly, the object to which another person is oriented is a basic datum for the *O* of that person.[4] One of the most important sources of this information is the line of regard of the eyes, the glance of the person. Visual 'looking' is often equivalent to paying attention, or being oriented toward an object or event. This orientation is the directional aspect of the face, which may at the same time have expressive properties as well. To know that the person is looking is to know that he is in contact with his environment, and that, therefore, he can be expected to act reasonably.

The role of looking is clearly a large one in the facial configuration, and it is surprising that so little work has been done on the topic. Heider's discussion makes a significant theoretical contribution,[5] and Exline has provided an interesting study of looking behavior in a small-group situation.[6] It remained for Gibson and Pick to

---

* Received for publication March 10, 1966. The experimental work was done at the University of Maryland. Experiment I supported by the University of Maryland General Research Board; Experiments II and III supported in part by Grant MH06644 from the National Institute of Mental Health. Stuart Schaffer aided in collecting and analyzing part of the data.

[1] Harold Schlosberg, Three dimensions of emotion, *Psychol. Rev.*, 61 1954, 81–88.

[2] A. E. Michotte, The emotions regarded as functional connections, in M. L. Reymert (ed.), *Feelings and Emotions*, 1950, 114–126.

[3] Fritz Heider, *The Psychology of Interpersonal Relations*, 1958, 37-44, 79-124.

[4] M. G. Cline, The influence of social context on the perception of faces, *J. Person.*, 25, 1956, 142–158; N. H. Frijda, Facial expression and situational cues, *J. abnorm. soc. Psychol.*, 57, 1958, 149–154.

[5] Heider, *op. cit.*, 37-44, 79-124.

[6] R. V. Exline, Explorations in the process of person perception: visual interaction in relation to competition, sex and need for affiliation, *J. Person.*, 31, 1963, 1–20.

analyze the stimulus for the perception of looking behavior in another person.[7] We shall be concerned here with further analysis of this stimulus.

Gibson and Pick selected a single case of the perception of where a person is looking: the perception of being looked at. Consequently, their experiment involved an *S* looking at a looker (*L*) and judging whether or not he, the *S*, was being looked at. They report that the acuity shown by their *S*s fell within the normal range. In considering the nature of the distal stimulus for this percept, they concluded that it was not merely the centering of the iris in the visible sclera, but also the position of the eyes in the head and their relation to the position of the head in space. Each eye-position, when combined with a given head-position, specified the direction in which the eyes were looking. This conclusion was based on the results they obtained when the *L*'s head was turned to the right or left rather than straight ahead. Under these conditions, *S*'s acuity was not impaired, although there was a constant error in the direction of turn of *L*'s head.

The present investigation attempts (1) to replicate these results and (2) to consider acuity for targets other than those located beween *S*'s eyes. We have also measured constant errors for these judgments. Further, we shall consider the Gibson-Pick hypothesis that the proximal stimulus for the perception of direction of gaze is the invariant resultant of eye-head positions. Both the position of *L*'s head as well as the direction of his eyes are varied in these experiments.

METHOD

An apparatus was constructed which provided *L* with a view of a variety of targets and *S* with a frontal view of *L*'s eyes and face. This was accomplished with a half-silvered mirror, as in Fig. 1. The target on which *L* fixated was 122 cm. from *L*. This was the same as the apparent distance of *L* to *S* (*i.e.* the sum of the distances of each to the mirror). Under these conditions, *L* always fixated on targets at the same distance from him as was *S*.

The target-board (Fig. 2a) contained a series of targets so arranged in cross-hair fashion that the point of intersection, the center dot, was directly in front of *L* and at the same level as his eyes. The dots were so arranged that they represented 2°, 8°, and 12° of angular rotation of *L*'s eyes, upward, downward, right, and left.

The transparent response-board (Fig. 2b) contained 65 dots including 13 with the same coördinates as the target-dots.

The response- and target-boards were optically so aligned that the common dots on each board would overlap if projected onto a common plane. Thus, with both *S*'s and *L*'s heads held in position by their respective head-rests, and with *L* fixating on the center target, her line of regard passed through the midpoint between *S*'s eyes. For each target fixated by *L*, there was a corresponding dot on the response-board through which *L*'s line of regard passed.

The *L* was a woman with optically balanced eyes. She was instructed to hold her fixation until told to shift to a new target and to assume a passive facial expression.

---

[7] J. J. Gibson and A. D. Pick, Perception of another person's looking behavior, this JOURNAL, 76, 1963, 386–394.

Each *S* was told that the task was to judge where *L* was looking. If *L* was looking

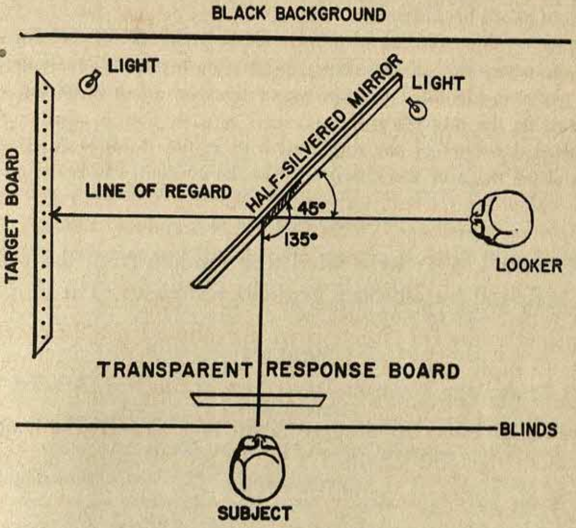


FIG. 1. PLAN OF APPARATUS

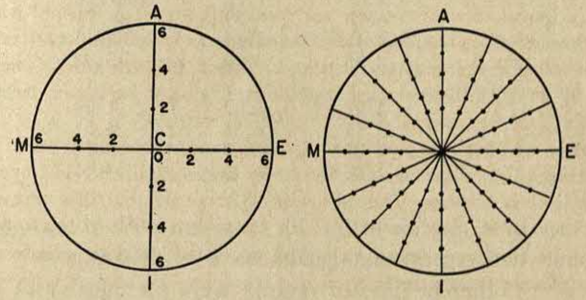


FIG. 2A. TARGET BOARD FIG. 2B. RESPONSE-BOARD

FIG. 2. APPEARANCE OF THE TARGET AND RESPONSE BOARDS TO *L* AND *S*, RESPECTIVELY

at him (in his eyes) he was so to report. Otherwise, he was to indicate through which dot on the response-board *L*'s line of regard passed. *S* was allowed to put his finger on the appropriate dot of the response-board. He could remove his head from the head-rest for a moment to check his judgment, but all judgments were made from the position of the head-rest.

Five practice trials were run and then *L* fixated each of 13 targets in a random sequence. The series continued until each target had been fixated 20 times, hence *S* made 260 judgments in all. Only stationary positions of the eyes were presented. Each view of *L* was presented for approximately 10 sec., which was always long enough for *S* to record his judgment.

Each judgment was scored for its errors in cm. Both vertical and horizontal components of the error were scored. Each judgment was scored on both dimensions, and each score was identified as positive or negative relative to the vertical and horizontal axes of the target-board.

The standard deviation of the distribution of errors is taken as the estimate of threshold, and the mean of the distribution as the constant error.

EXPERIMENT I

In Experiment I, 5 *S*s were selected at random from the pool of University of Maryland introductory psychology students. The only restriction was that they have approximately 20/20 vision with correction. Each *S* was run under two experimental conditions. First, 260 exposures of *L* (20 exposures of 13 different lines of regard) were presented with the lower half of *L*'s face covered by a piece of black cardboard. After an appropriate rest for both *L* and *S*, another 260 trials were run with *L*'s full face exposed.

TABLE I

CONSTANT ERRORS AND STANDARD DEVIATIONS, IN CM., FOR SELECTED TARGETS: EXPERIMENT I

(Positive vertical errors represent error above target; positive horizontal errors represent error to the right of target.)

Target	Vertical Error		Horizontal Error	
	CE	*SD*	*CE*	*SD*
Center	.56	2.70	.14	1.55
Vertical Displacement:				
A6 (Up 12°)	3.68	7.52	−1.92	5.58
A4 (Up 8°)	6.18	8.59	−.71	4.07
I4 (Down 8°)	−3.41	7.86	2.44	4.32
I6 (Down 12°)	−8.36	8.40	3.48	6.07
Horizontal Displacement:				
E6 (Right 12°)	4.07	5.62	1.09	6.49
E4 (Right 8°)	3.71	6.36	4.57	6.07
M4 (Left 8°)	4.87	6.37	3.55	5.92
M6 (Left 12°)	5.68	6.15	1.57	8.26

Results. A comparison of the data collected under the two main experimental conditions revealed no significant differences in *SD*s or *CE*s. The following data refer, therefore, to the full face condition alone.

Table I summarizes the *SD*s and *CE*s for a few representative targets.[8]

---

[8] Data for the remaining targets, essentially the same as those reported here, are available on request from the author.

These are displaced 17 and 25 cm. (8° and 12° angular displacement) to the right, left, above, and below the center target, as well as the center target itself.

Judgments of gaze at center target. Consider first the data of the present experiment most comparable to the data of Gibson and Pick. These are the horizontal components of the judgments of fixation on the center target (corresponding to the midpoint between *S*'s eyes). The horizontal threshold for this glance (*i.e.* the *SD* of the horizontal errors) is 1.55 cm. Translated into angular measure 1.55 cm. at 122 cm. subtends an angle of 0.75°. This represents the deviation in the eyes of *L* that *S* can just detect. On the assumption that the distance from the point of rotation of the eye to the anterior surface of the cornea is 13.5 mm.,[9] this angular rotation can be translated into a linear displacement of the iris of *L* of 0.18 mm. In other words, at 122 cm., a displacement of the iris of 0.18 mm., is just noticeable to *S* when the *L* is fixating on a target at the midpoint between *S*'s eyes.

If we assume that the threshold reported by Gibson and Pick at 200 cm. (calculated by the author to be 0.64 cm.) can be extrapolated in direct proportion, then their value is 0.39 mm., somewhat greater than our 0.18 mm. These distances represent 1.10′ and 0.51′ in terms of visual angle, respectively. Thus, while the present measure is slightly better than standard acuity, there is little reason to doubt the essential comparability of these *DL*s. Judgments of the line of regard of a looker which passes through the eyes of *S* can be made with very good accuracy at these distances.

Consider now the vertical component of the judgments of line of regard directed into *S*'s eyes (Table I). The vertical threshold is 2.70 cm. which can be translated into an angular rotation of 1.25° and a linear displacement of the iris of 0.3 mm. A one-tailed *t*-test of differences between correlated variances reveals the vertical threshold to be significantly larger than the horizontal threshold ($t = 6.28$, 98 *df*).

Thus, at 4 feet, a displacement of the *L*'s line of regard to a point about half an inch to the right or left of the *S*'s eyes is just noticeably different than being looked at, whereas a displacement of *L*'s eyes to a target about an inch above or below *S*'s eyes, to the middle of the forehead or nose, is required in order to be just noticeable. It would be instructive to know if there are any further psychological differences associated with these figures.

(2) Judgments of gazes at off-center targets. The *DL*s for these judgments range from 4.07 cm. to 8.59 cm. with a mean at 6.61 cm. All are larger than the *DL* for judgments of the center target gaze (one tailed sign test, $p < 0.03$). There is also an interaction between the nature of the displacement of the line of regard and the relative strength of the

---

[9] See Gullstrand's model in O. Stuhlman, Jr., *Introduction to Biophysics,* 1943, 102.

two componants of the error. If we sum the vertical variances and the horizontal variances for the vertical and horizontal targets separately, the interaction can be made apparent. It is represented graphically in Fig. 3. When the line of regard is displaced up or down, the vertical component of the error is greater than the horizontal component. On the other hand, when the line of regard is displaced to the right or left, it tends to be judged with greater error in the horizontal component than the vertical component.

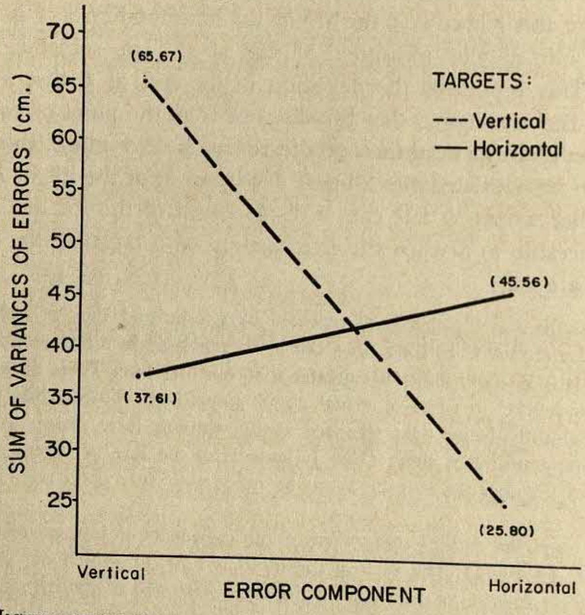


FIG. 3. INTERACTION BETWEEN MAGNITUDE OF ERROR AND DISPLACEMENT OF TARGET

Conclusion. One possible explanation of these results is that the stimulus with which we are dealing is the position of the iris relative to the visible sclera. If the iris is vertically centered in the sclera, but horizontally displaced, then judgments of horizontal position should be more difficult than vertical position. This might also contribute to the relative ease of the center dot, since the *L*'s iris for these targets is centered in both the vertical and horizontal dimensions.

Although this explanation is a powerful one, it cannot be complete. If the stimulus with which we are dealing is the position of the iris in the sclera, then the judgments of all gazes which have, for example, horizontally centered irises, should have equally small horizontal errors. Judgments of gazes with vertically centered irises should have small vertical errors. The gaze directed at the center target has an iris centered both vertically and horizontally, but both components of the error for this gaze are significantly smaller than the corresponding components for any other gaze whether centered or not. The line of regard which produces the judgment: "She is looking at me" cannot be accounted for exclusively in terms of the centering of the iris in the sclera.

Gibson and Pick examined this issue by having *L* turn her head to the right or left relative to *S* while gazing into *S*'s eyes.[10] They found a *CE* under these conditions, but they reported no significant change in acuity compared to the head-straight position. Although it is surprising that a sidewise glance-in-the-eye can be judged with the same degree of acuity as a full face glance-in-the-eye, the Gibson and Pick data tend to support the notion of the present experiment, that the look in the eye is a special case of the more general situation of looking. This will be explored below.

With respect to the *CE* for these targets, there is a clear upward *CE* for all targets at the level of the eyes or higher, and a downward *CE* for targets in the region of the mouth. At the same time, there is a *CE* to the right for all targets at the level of the eyes and below and a *CE* to the left for the targets above the level of the eyes.

EXPERIMENT II

This experiment is designed to study the effect of *L*'s head turned away from the full face position on accuracy and constant errors.

Procedure. The apparatus was modified slightly to facilitate the experimental procedure. The change involved replacing the painted targets on the target-board, with small neon lights wired to an interval-timer and a selector switch. *E* could select the target on which *L* was to fixate by turning the switch to the appropriate position. A neon bulb at that position then flashed regularly at half-second intervals.

The procedure in this experiment was also modified. We were no longer interested in collecting data on such a large number of lines of regard, or in comparing the relative acuity of the vertical and horizontal components of the judgments. Thus, targets displaced vertically and responses in the vertical direction were eliminated. Five horizontal targets were used (4° and 10° to the right and left, plus the center target) together with a number of targets introduced at other locations to break up response sets. The *S*s were seven undergraduates randomly selected from the same pool as that from which the *S*s in Experiment I were drawn.

Results for Experiment IIA. (Experiment IIA was a replication of Experiment I on the new apparatus.) The horizontal threshold for the center target was raised slightly to 2.85 cm. although some of the thresholds for the other lines of regard were reduced, (Table II). The patterns of these thresholds and the *CE*s replicate the first experiment.

Results for Experiment IIB. In Experiment IIB, *L*'s head was held 30° to the right by a head clamp. A constant error in the direction of the turn of the head is present in the resulting judgments, given in the right hand columns of Table II, except for the case of extreme discrepancy between head and eye orientation. With the head turned 30° to the right and the eyes aimed 10° to the left of the center target, there is no *CE*. The *CE*s are generally large.

---

[10] Gibson and Pick, *op. cit.*, 389, 392.

In this experiment, the accuracy in judging the line of regard which passes through the eyes of *S* is significantly lower than accuracy in judging that line of regard with the head straight. The *DL* for this stimulus is 5.29 cm., which is significantly greater than the *DL*s in Experiments IIA and I. These data do not confirm the finding of Gibson and Pick, who report that there is no difference in acuity for judging the line of regard which passes through the eyes when *L*'s head is turned or straight.

TABLE II

CONSTANT ERRORS AND STANDARD DEVIATIONS, IN CM., FOR EXPERIMENT IIA AND IIB

Target	Experiment IIA (Full face)		Experiment IIB (Head turned right)	
	CE	*SD*	*CE*	*SD*
Left 10°	−.32	5.78	−.54	5.09
Left 4°	.63	4.94	2.64	6.60
Center	.23	2.85	3.00	5.29
Right 4°	2.58	3.54	4.78	4.62
Right 10°	3.33	5.19	5.26	5.94

Conclusions. The effect of a divergence between the orientation of the head and eyes is to increase the *CE*s and lower acuity. This is true for almost every line of regard, and, in the sample of lines of regard in the present study, this conclusion is independent of the position of the iris to the right or to the left for any given line of regard. For example, when aiming at a target to the right with the head straight, the iris is on the right side. The identical line of regard with the head turned 30° to the right has the iris on the left side. The *CE*s and *DL*s are the same in these two situations. So far, the only instance examined in which there is no divergence between the direction of the eyes and of the head is the special case of the head and eyes aimed at the eyes of *S*. It is now necessary to consider non-divergent orientations of the eyes and head for lines of regard not directed into the eyes of the *S*.

EXPERIMENT III

Experiment III was conducted with *L* looking at each of the targets used in the previous experiment, but with her head turned parallel to her eyes for each line of regard. In the earlier experimental conditions, the orientation of the head and eyes were diverged, except for the center target in Experiments I and IIA.

The *S*s were seven new undergraduates from the same subject pool.

Results. Table III indicates that judgments of the center target in Experiment III are equal in accuracy to the judgments of the same target in Experiment IIa. This was expected.

An analysis of variance of the *SD*s of judgments of the five lines of regard across the three experimental conditions (Experiments IIA, IIB, and III)[11] reveals that the only significant main effect is between lines of regard. The Duncan New Multiple Range Test shows this to be produced by the significantly lower thresholds for the regard of the center target ($p = 0.001$).[12] There are no significant differences between thresholds for the non-center targets.

The *CE*s are, however, quite small for all lines of regard in this experiment, with a slight tendency in the direction of the turn of the head and eyes. Thus, when the head and eyes of *L* are turned in the direction of the targets to the left, the *CE*s are 0.21 and 1.40 to the left, and when the head and eyes are turned to the right the *CE*s are 0.32 and 0.83 to the right. All *CE*s in Experiment III are significantly smaller than those in Experiment IIb, except for the left 10° gaze (extreme discrepancy). The effect of maintaining the head and eyes parallel to each other is to reduce the constant error produced by turning the head 30° to the right. If the discrepancy between the head and eyes involves the turning of the head away from the full face position with respect to the *S*, the constant error tends to be in the direction of the turn of *L*'s head. If the discrepancy involves only the turning of the eyes (Experiment I), then the constant error tends to be to *S*'s right. If there is no discrepancy between *L*'s head and eyes, then the constant errors are at a minimum. If there is no discrepancy and both the head and eyes are turned away from the *S* then there is a slight constant error in the direction of the turn of the head. The primary effect of a discrepancy between head and eyes is to reduce the acuity of judgments of being looked at.

TABLE III

CONSTANT ERRORS AND STANDARD DEVIATIONS, IN CM., WITH HEAD AND EYES PARALLEL: EXPERIMENT III

Target	*CE*	*SD*
Left 10°	−.21	5.81
Left 4°	−1.40	4.02
Center	−.61	2.83
Right 4°	.32	5.88
Right 10°	.83	6.75

---

[11] The Hartley *F max.* test indicates that the variance is not significantly heterogeneous at the 0.05 level. See H. O. Hartley, The maximal F-ratio as a short-cut test for heterogeneity of variance, *Biometrika,* 37, 1950 308–312.

[12] A. L. Edwards, *Experimental Design in Psychological Research,* 1960, 136.

Discussion. It is apparent that there is a good deal of evidence in support of the Gibson and Pick hypothesis that there is total stimulus, combining the direction of the eyes and the orientation of the head, that defines a direction of regard. As now stated, however, this hypothesis is incomplete. Head-position and eye-position interact to produce a perceived direction which falls between these two positions. If both the head and eyes are oriented in the same direction then there is almost no constant error to *S*'s judgments, and this holds for a wide range of direction. If, however, the head alone remains straight ahead while the eyes vary in their orientation, there will be constant errors characteristic of the direction of the eyes.

In the vertical dimension we have found large constant errors which cannot be related to head position. Acuity is reduced when either the head or the eyes are moved out of the straight-ahead position with respect to *S*. This last finding disagrees with the results of Gibson and Pick. Our results suggest that the combination of head and eyes directed straight at *S* is a special stimulus indeed. As a matter of fact, for most head-positions, the line of regard directed into *S*'s eyes is discriminated with greater accuracy than all other lines of regard. The mutual glance is a unique experience and there are unique judgments coördinated with it.

The complex interaction between the displacement of *L*'s eyes and the magnitude of *S*'s errors suggests a systematic mode of perceiving the glances of another. Such a perceptual posture may imply that different kinds of meanings are carried by the different glances. Since sensitivity is generally greater for horizontal displacements than for vertical displacements of the eyes, certain interpersonal meanings might very well be more typically carried by the horizontal displacements. The shifty-eyed, embarrassed, or nervous looker may be prone to break visual contact by a horizontal eye-movement while the thoughtful looker may look upward in order to think more clearly and the chastised looker may cast his gaze down to signify repentence. These are by no means firm predictions, but the attempt to coördinate a wide range of interpersonal relations to specific eye-head stimulus-patterns could be fruitful.

SUMMARY

S was presented with a mirror image of a person looking at a number of targets scattered about his (*S*'s) head, including one target located between his eyes. *S*'s task was to determine where the *L* was looking. Thresholds and constant errors for these judgments were determined and compared for the various targets under a number of experimental conditions. Accuracy for being looked at is quite high, and accuracy for other lines of regard are somewhat lower. Patterns of constant errors are apparent. By comparing judgments of the lines of regard of a *L* whose head was, first, always directly facing *S*; second, always turned to the right; and, finally, turned in each case to look in the direction of the line of regard, the Gibson and Pick hypothesis that head position contributes to the judgment of the direction in which the eyes are looking was generally supported. The single condition where the glance of *L* is directed at a point between the eyes of *S* may play a unique role in these judgments.

CROSS-MODAL JUDGMENTS OF SMALL HOLES

By S. M. ANSTIS and C. M. LOIZOS, University of Bristol, England

Although one expects judgments of size to vary to some extent from person to person, one might expect the successive judgments of an individual to be fairly consistent, given constant viewing conditions, any variation amounting to no more than random error. We have, however, more than one modality by which we may judge size; the question is now whether there is exact correspondence between any two of these within a single individual. In other words, do things look the same size as they feel? In the following experiment, the subjective sizes of small holes were compared by touch and sight.

Berkeley (1709) saw clearly the difficulties in investigating such a problem.[1] He wrote, "The visible object still changing as you approach or recede from the tangible object, it hath no fixed or determinate greatness. Whenever, therefore, we speak of the magnitude of any thing . . . we must mean the tangible magnitude; otherwise there can be nothing free from ambiguity spoke of it." If this is true, it should be impossible to make reliable cross-modal comparisons, a task which several experimenters have attempted.

In 1886, Jastrow investigated the perception of short lengths (5 to 120 mm.) by disparate senses.[2] Eye, hand and arm were presented with varying distances which were then reproduced by the same, or one of the other two sense organs. Jastrow found that the length of a felt object could be matched visually, and vice versa, but that systematic illusions of size occurred. When the 'receiving' and the 'expressing' sense are the same, the "error is small and the process relatively easy." When, however, different senses acted as receiver and expresser, results varied considerably. Regardless of which was the receiving sense, lengths expressed by the hand were overestimated and by the eye underestimated. When the arm was the means of expression, lengths presented to the eye were overestimated, but lengths presented to the hand were underestimated. In every case, the extent of the error decreased as the length to be judged increased. Jastrow concluded that it was the expressing sense that determined the directions and extent of the error. Moreover, all the relations he established were reflexive, in that "if reproducing one sense by another results in an exaggeration (or underestimation), then reproducing the second sense by the first will result in an

---

* Received for publication March 28, 1966. The first author was supported by a research grant at the University of Cambridge under U. S. Air Force Contract AF EOARDC 63–93, supervised by R. L. Gregory. The second author was supported by an SRC research studentship at University College, London.

[1] George Berkeley, An essay towards a new theory of vision, Sections 49, 55, 1709, 34-36.

[2] Joseph Jastrow, The perception of space by disparate senses, *Mind*, 11, 1886, 539-554.

underestimation (or exaggeration) to about the same extent." Nevertheless, receiving and expressing by different senses was a task both "inaccurate and difficult." Jastrow felt therefore that "the connection between the senses seems to be a loose one," and concluded that the order of differential sensibility of the three senses was sight, span (hand), and motion (arm).

Unlike Jastrow, Kelvin (1954) found no significant difference between ipse-modal and cross-modal judgments, using disk sizes rather than varying short lengths.[3] However, Kelvin and Mulik (1958) then studied discrimination of length by sight and touch and on the basis of their results concluded that ". . . it is doubtful if cross-modal matching is ever more than an experimental artefact."[4] Yet, both Jastrow and Kelvin agree that it is the expressing rather than the receiving sense that determines the nature of the results.

In the present study, cross-modal comparisons of small holes are made in an attempt to determine the existence and the extent of cross-modal illusions in the judgment of size. It is, of course, impossible to ask which of the subjective sizes is nearer to the object's real, or physical size, since there is no direct way of comparing objective with subjective hole-size.

Method. Six experiments were carried out on each of the 9 *Ss* in each of which the *Ss* made cross-modal comparisons of the subjective sizes of small holes using a simultaneous matching method. Four sizes of standard hole were matched with a comparison range of 32 hole-sizes.

Subjects. The *Ss* used were 9 graduate students, 6 men and 3 women, aged 22-30 yr.

Apparatus. The sizes of the four standard holes were 0.125, 0.25, 0.375, and 0.4375 in. Each was reamed in the center of a 2 in. square of light brown plastic, 0.125 in. thick. The comparison-range of holes was reamed around the circumference of a 5.5-in. diameter disk of the same material. They ranged in size from 0.0625 to 0.5 in. in graduated steps of 0.0156 in. (1/64th). The disk was spring mounted on a spindle so that the subject could, for example, rotate it with his thumb while feeling the holes with his finger.

Procedure. Each of the four sense-organs used (eye, tongue, right index finger, left little finger) was compared with every other, giving six comparisons.

Standard stimulus	Comparison
(1.) Tongue	Right index finger
(2.) Tongue	Left little finger
(3.) Eye	Right index finger
(4.) Eye	Left little finger
(5.) Tongue	Eyes
(6.) Left little finger	Right index finger

Two finger-sizes were used to determine whether finger-size alone would affect the results. The left-hand was chosen since in Experiment 6, manipulating both the standard and the comparison-ranges of holes with one hand was found to be awkward.

---

[3] R. P. Kelvin, Discrimination of size by sight and touch, *Quart. J. exp. Psychol.* 6, 1954, 23-34.

[4] R. P. Kelvin, Discrimination of size by sight and touch. *Quart. J. exp. Psychol.,* 10, 1958, 187-192.

The procedure was similar throughout. In Experiment 1, *S* sat with his eyes closed and palpated with his tongue the standard stimulus-hole in the plastic square, which was held to his lips with the left hand. At the same time, he ran his right index finger around the holes in the comparison-disk, until he selected the hole which felt the same apparent size as the standard hole. In Experiment 2, he held the standard square to his lips while running his left little finger round the comparison-disk. In Experiments 3 and 4, the standard square was placed on a table against a black background, at right angles to the line of sight, and viewed binocularly from a distance of 18 in. *S* felt the comparison-disk with his little or index finger under a screen. In Experiment 5, he felt the standard square with his tongue while viewing the comparison-disk, which was mounted in the vertical frontal-parallel plane 18 in. from his eyes. He then pointed with a rod to the hole selected as a match. In Experiment 6, he felt the standard hole with his left little finger and the comparison-disk with his right index finger.

Each standard hole was presented eight times during each experiment, in random order, and the comparison-disk was rotated through a random angle between trials, to counteract positional habits. The rod used to select the appropriate visual hole on the comparison-disk in Experiment 5 was a 12-in. piece of O-BA studding. In an attempt to reduce possible visual cues from the thickness of the rod, *S* was instructed to make his selection before raising the rod into his field of vision.

The experiments were conducted consecutively with each *S*, with a short rest between successive ones. The order was permuted between *S*s to reduce practice effects, but the experiments involving vision (Experiments 3 and 5) were put at the end as it was thought they might give the *S*s too much information about the experimental arrangements. The *S*s were allowed to see the comparison-disk before the experiments but not the four standard squares. They were not allowed to see any holes during the course of the experiments, except in the visual conditions.

Results and discussion. Results are shown on Table I and represented graphically in Fig. 1. Inspection of Fig. 1 shows that holes presented to the tongue and matched by either of the two fingers are overestimated (Fig. 1.1, 1.2). The holes selected as a match vary between 99 and 162% of the true hole-size, with a mean of approximately 122%.

The graphs for tongue vs. right index finger (Fig. 1.1) and for tongue vs. left little finger (Fig. 1.2) are similar, indicating that it makes no difference which finger is used.

Holes presented to the eyes and reproduced by the fingers are overestimated (Figs 1.3, 1.4). The holes selected as a match vary from about 98 to 150% of the true hole size, with a mean of approximately 117%. The graphs for eye vs. right index finger (Fig. 1.3) and for eye vs. left little finger (Fig. 1.4) are similar, again indicating that finger-size does not affect the results.

In all these conditions, the error decreases as the size of the holes to be judged increases, thus confirming Jastrow's findings. Moreover, all four graphs discussed so far (Figs. 1.1, 1.2, 1.3, 1.4) are similar, showing that the eyes and the tongue overestimate holes in a similar way. This is supported by Fig. 1.5, in which it is seen that holes presented to the tongue and reproduced by the eyes are matched fairly accurately, as are holes presented to the left little finger and reproduced by the right index finger (Fig. 1.6). The similarity of the two graphs (Figs. 1.5, 1.6) suggests that tongue vs. eye matching and right index vs. left little finger matching are about equally accurate.

TABLE I

REGRESSION-FUNCTIONS OF THE SIZE-MATCHES IN FIG.1

Observed regression-function		S.D.	V.R.		Predicted regression-function	V.R
(1)	$I=.772T+.102$	.051	1.049	IET	$I=.735T+.104$	1.036
				ILT	$I=.702T+.122$	1.063
(2)	$L=.770T+.110$	.049	1.019	LET	$L=.721T+.116$	1.053
				LIT	$L=.846T+.088$	1.043
(3)	$I=.804E+.085$	.034	1.074	ITE	$I=.844E+.082$	1.097
				ILE	$I=.719E+.111$	1.096
(4)	$L=.789E+.098$	.030	1.024	LTE	$L=.842E+.090$	1.229
				LIE	$L=.881E+.069$	1.130
(5)	$E=.914T+.024$	.073	1.018	EIT	$E=.960T+.021$	1.045
				ELT	$E=.975T+.016$	1.050
(6)	$I=.912L+.022$	.033	1.050	ITL	$I=1.002L-.009$	1.141
				IEL	$I=1.018L-.015$	1.219

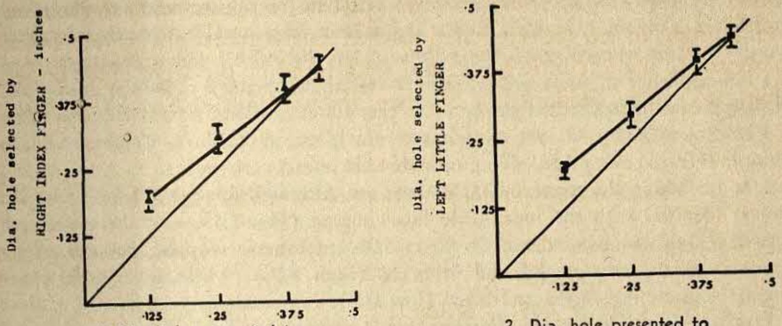
I = Right index finger
L = Left little finger
E = Eye
T = Tongue

The observed regression-functions, on the left of the table, were fitted to the data points by least squares. The predicted regression-functions, on the right of the table, were obtained by combining pairs of observed regression-functions. Thus, the observed functions

$I=.804E+.085$ (left column, row 3)
and $E=.914T+.024$ (left column, row 5)
yield the predicted function
$I=.804\,(.914T+.024)+.085$
Multiplying out,
IET $I=.735T+.104$ (right column, row 1)

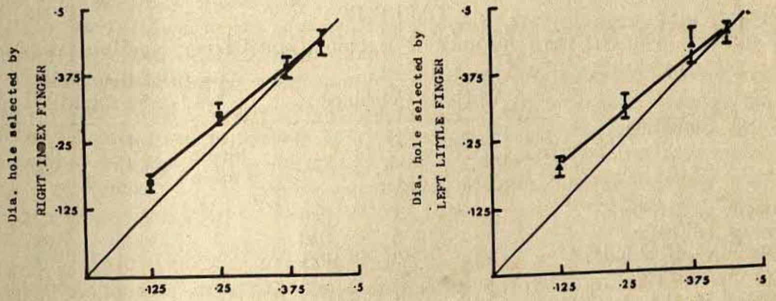
In summary, holes are judged larger when presented to tongue or eyes than when presented to fingers. They appear approximately the same size to the right index finger as to the left little finger. Errors decrease as the hole-sizes increase; they are greatest for the smallest holes (0.125 in.), becoming minimal at a hole-size of 0.4375 in. Statistical treatment of the results reveals the following:

Linearity. Each experiment gives a set of four points. A regression-line was fitted to each set of four points by the method of least squares, and tested for linearity



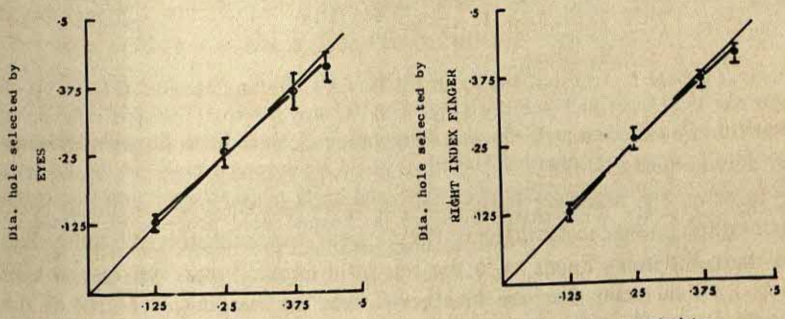
1 Dia. hole presented to TONGUE—inches

2 Dia. hole presented to TONGUE



3 Dia. hole presented to EYES

4 Dia. hole presented to EYES



5 Dia. hole presented to TONGUE

6 Dia. hole presented to LEFT LITTLE FINGER

FIG. 1. CROSS-MODAL SIZE-MATCHES

by Fisher's variance ratio test. Each set could be fitted by a linear equation (regression-function) of the form $y = ax + b$. These regression-functions are given in Table I, Column 1. In each case, a is less than unity, and b is positive. In other words, the line on each graph has a slope of less than unity, and is displaced above the indicated line of unity slope. Table I also shows that $a \propto 1/b$: the greater the displacement, the smaller the slope.

Variance. *SD*s for all six experiments are given in Table I, Column 2. Each point is averaged over all *S*s, giving intra- but not inter-*S* variance.

*SD*s lie within the range 0.033 to 0.051 in. Although the eye-tongue means are almost identical with the index-little finger means (Fig. 1.5, 1.6), the eye-tongue *SD*s are high and the index-little finger *SD*s are low. It may be that eye-tongue comparisons are cross-modal and 'difficult' (high *SD*s), while index-little finger comparisons are ipse-modal and 'easy' (low *SD*s).

Organ responsible for the illusions. Fig. 1 shows that a size-illusion exists for comparisons between tongue and finger (Figs. 1.1, 1.2), and between eye and finger (Figs. 1.3, 1.4) but hardly at all between eye and tongue (Fig. 1.5) or between one finger and another (Fig. 1.6). It may be that illusions only occur when the fingers take part. On this interpretation, the lack of illusions for comparisons between one finger and another would be a special case, on the grounds that such judgments are more 'compatible' and hence more accurate than true cross-modal judgments. What characteristic of the fingers could generate the illusions? The most likely candidate is lack of discrimination. The fingers are working near the limits of their resolution in judging holes of 0.125-in. diam. The fingers have poorer spatial acuity than the tongue and eye: the two-point threshold of the fingers is about 0.1 in., compared with 0.04 in. for the tongue. The eye's resolving power is, of course, orders better than this.

TABLE II

COMPARISON OF *SD* WITH AMOUNT OF ILLUSION, FOR DIFFERENT MODALITIES

Condition	All Holes Combined			Smallest Hole Size (0.125 in. diam.)		
	Size-illusion rank order	*SD* rank order	*SD*	Size-illusion	*SD* Rank order	*SD*
Tongue vs. left little	1	3	.049	1	2	.023
Eye vs. left little	2	6	.030	2	5	.017
Tongue vs. right index	3	2	.051	3	1	.024
Eye vs. right index	4	4	.034	4	4	.019
Tongue vs. eye	5	1	.073	5	3	.022
Left little vs. right index	6	5	.033	6	6	.017
Correlation between size-illusion and *SD*		rho= −0.086 (NS)			rho= +0.486 (NS)	

The data do not, however, support this interpretation too strongly. If size-illusions were due to poor spatial acuity in the fingers, then there should be a correlation between amount of size-illusion, and poor discrimination as measured by *SD*. This correlation should be greatest for small holes, where the fingers are at their greatest disadvantage. In fact, Table II shows that there is virtually no correlation between the amount of illusion and the *SD*, for all hole sizes combined (rho $= -0.086$: not significant). For the smallest hole size (0.125-in. diam.), there is a correlation (rho $= +0.486$), but it is not significant. So we cannot conclude definitely that size-illusions are due to poor discrimination by the fingers.

Transitivity. The results of any two experiments can be used to predict a third. For example, if it is observed that $I = f(E)$ and that $E = g(T)$, then it is possible to predict that $I = \mathrm{f.g(T)}$. If the prediction is confirmed by observation the functions may be called transitive.

The observed and predicted functions are shown in Table I. In Column 1, the observed functions from the six experiments are combined in pairs to yield the 12 predicted functions of Column 4. For example, in Column 1,

equations (3) $I = .804\mathrm{E} + .085$ and

(5) $E = .914 + .024$

yield the predicted equation in Column 4, Row 1,

IET $I = .735T + .104$

The letters *IET* indicate that this equation is derived from equations $I = f(E)$ and $E = g(T)$, with *E* as an intermediate term.

The predicted functions were compared, point by point, with the observed functions, using Fisher's variance ratio-test. The mean variance ratio, averaged over four points, is given in Column 3. In each case, the low variance ratio found is consistent with the null hypothesis of transitivity at better than 5% level of confidence ($F = 1.47$ for the point by point 71 *df*, and $F = 1.26$ for the slope (mean of four points) 286 *df*).

We conclude that all the size-judgment functions studied are transitive. This finding may be compared with Jastrow's findings of reflexivity; *i.e.*, that when $\mathrm{Y = f.X}$ it is found experimentally that $\mathrm{X = f^{-1}Y}$. Thus the answer to our original question, do things look the same as they feel,[5] is that they do not. Holes of less than half an inch look larger than they feel, and feel larger to the tongue than they do to the finger (a fact verifiable by anyone with an impending visit to the dentist).

Should one expect there to be correspondence in size-judgments between different sense modalities? It is suggested that the answer to this question is 'no'; different sense-organs have evolved to provide different types of information about the environment. By the very nature of their differing and specialized equipment, no two of them will provide identical answers to the same question; nor were they intended to do so. As Berkeley puts it, "If we take a close and accurate view of the matter, it must be acknowledged that we never see and feel one and the same object. That which is seen is one thing, and that which is felt is another."

---

[5] S. M. Anstis, Apparent size of holes felt with the tongue, *Nature* 203, 1964, 792-793.

SUMMARY

Cross-modal judgments of small holes were made in an attempt to determine the existence, if any, and the extent of cross-modal illusions in the judgment of size. Each of the four sense-organs used (eye, tongue, right index finger, left little finger) was compared with every other, giving six comparisons in all. Each of 9 *Ss* was tested on all six comparisons, using a simultaneous matching method. Four sizes of standard hole were matched with a comparison-range of 32 hole sizes.

Results showed that holes are judged larger when presented to tongue or eyes than when presented to fingers. They appear approximately the same size to the tongue as to the eyes; and approximately the same size to the right index finger as to the left little finger. The extent of the error decreases as the size of the hole to be judged increases: they are greatest for the smallest holes (0.125 in.) and become minimal at a hole-size of 0.4375 in. The reasons for this discrepancy are discussed.

RATIO-ESTIMATIONS OF LOUDNESS-INTERVALS

By JACOB BECK, Harvard University, and WILLIAM A. SHAW, University of Pennsylvania

Two loudness-scales have been published: (a) the sone scale,[1] and (b) the lambda scale.[2] These two scales, based on different scaling procedures, are quite divers. The sone scale, based on direct estimates of the apparent ratios of loudnesses, is a power function of sound-intensity with an exponent of 0.3. The lambda scale, based on both estimates of equal-appearing intervals and the apparent ratios of loudnesses, is much less steep and may be approximated by either a logarithmic function or a power function with an exponent of 0.13. Though the lambda scale, founded on both equisectional and fractionational judgments, is a hybrid scale, the form of the scale based primarily on equisectional data is in general agreement with both discriminability and interval scales.[3]

The present experiment deals with the adequacy of the sone and the lambda scales to predict ratio-estimations of loudness-intervals. Ordinarily, the apparent magnitude of loudness-ratios is studied by ratio-scaling methods, and the apparent magnitude of loudness-intervals by interval scaling-methods.[4] That is, the scalings of ratios and of intervals are coupled to two different scaling tasks: that of making ratio-estimations and that of partitioning a segment into equal intervals. By using a ratio-method to study the apparent magnitude of loudness-intervals, the present experiment unties the property scaled from the method of scaling.

Method. Two independent groups of *O*s made ratio-judgments of intervals selected from the sone and lambda scales. Twenty-eight *O*s made judgments of intervals chosen from the sone scale, and 29 made judgments of intervals chosen from

---

* Received for publication May 12, 1966. This research was supported by Grant G13942 from the National Science Foundation. The preparation of this report was supported by NSF Grant GB-1521.

[1] S. S. Stevens, The measurement of loudness, *J. acoust. Soc. Amer.*, 27, 1955, 815–829.

[2] W. R. Garner, A technique and a scale for loudness measurement, *ibid.*, 26, 1954, 73–88.

[3] W. R. Garner, Advantages of the discriminability criterion for a loudness scale, *ibid.*, 30, 1958, 1005–1012; R. D. Luce and Ward Edwards, The derivation of subjective scales from just noticeable differences, *Psychol. Rev.*, 65, 1958, 222–237, esp. 235.

[4] Stevens, Ratio scales, partition scales, confusion scales, in Harold Gulliksen and Samuel Messick (eds.), *Psychological Scaling: Theory and applications*, 1959, 49–66.

the lambda scale. The *O*s were undergraduates at the University of Pennsylvania who were paid to participate.

Each *O* was run individually in a soundproof room. The stimuli were produced by a Grayson-Stadler twin channel oscillator and delivered via a TDH-39 earphone to *O*'s right ear. The general procedure is best described by the instructions given.

Instructions. On each trial you will hear two pairs of tones delivered as follows: (1) beep–(2) beep——(3) beep–(4) beep. The first pair, (1) and (2), will be called the standard. This pair will remain the same throughout the experiment and will always occur first on every trial. You are arbitrarily to call the distance between these two loudnesses 100. Your task is to judge the distance between the second pair of tones, (3) and (4), relative to that of the standard. That is, the number assigned to the distance between the second pair of loudnesses is to be proportional to the distance of the standard. For example, if the distance between the second pair of loudnesses seems to be about 5 times as much as that of the standard, you are to label the second pair 500. On the other hand, should the distance appear to be only 1/10 that of the standard, you are to label the second pair 10. Please give your judgments through the microphone as soon as you have heard both pairs. You will then proceed to the next trial.

The standard and comparison loudness-intervals consisted of two 1-sec. tones separated by ½ sec. The separation between the standard and comparison intervals was 2 sec. Before each stimulus-presentation, *E* adjusted the intensities of the comparison-loudnesses by means of decibel boxes. All timing was controlled electronically and the entire stimulus-sequence was set off automatically when *E* pressed a button.

The experimental plan was similar for both groups. Four intervals of 5, 10, 15, and 20 sones were used with Group I, and four intervals of 0.85, 1.70, 2.55, and 3.40 lambda units were used with Group II. The stimuli were 1000 cps. tones. Table I shows the intervals and their boundaries in decibels re 0.0002 dyne/cm^2 (*SPL*). The loudness-intervals were selected as follows: First, four successive 5-sone intervals were so chosen that the upper and lower boundaries of neighboring intervals should be close to one another. The four intervals covered a range of intensity from 30 to 86.6 db. *SPL*. The 10-, 15-, and 20-sone intervals were then so selected that they would begin at the lowest value chosen (30 db.) and end at the highest (86.6 db.). The lambda intervals were chosen to correspond as closely as possible to the sone intervals. The range of intensities was thus divided into four sectors. The four sectors are shown in Table I labeled lowest intervals, lower middle intervals, upper middle intervals, and highest intervals. The second 10 sone-interval (2 and 12 sones) and the second lambda interval (1.4 and 3.1 lambda units) had effectively the same boundaries (50 and 76 db.) and served as standards.

The *O*s made four judgments of each of the 16 comparison-intervals. Two judgments were made when both the standard- and comparison-intervals were presented in a loud-soft order, and two judgments were made when the intervals were presented in a soft-loud order. Because *O*s had difficulty when the loud-soft and soft-loud presentations were intermixed, the 64 judgments made by an *O* were divided into four series, each containing the 16 different intervals. The four series were presented alternately in a soft-loud and loud-soft order. The intervals were presented in a random order in each series. The median of the four estimations made by an *O* was taken as representative of his judgment.

Results. The questions posed are these: (1) Will intervals equal on the sone or the lambda scale yield similar ratio estimates? (2) Will *O*s' ratio-judgments be linear with the lambda or the sone scale?

Fig. 1 and 2 present the medians of *O*'s median estimates for Groups I and II. The curves in the lower half of Fig. 1 plot *O*s' judgments with respect to sone intervals. The curves in the upper half of Fig. 2 plot the judgments with respect to lambda-intervals. These results indicate that equal intervals on the lambda scale yield similar estimates, while equal intervals on the sone scale yield dissimilar estimates. The *O*s' median judgments for each of the sone intervals are consistently greater for the intervals chosen from the lower end of the scale than for intervals chosen from the higher end. Differences in *O*s' estimates of intervals chosen from the lambda-scale, however, show no such relationship. A Friedman rank-order analysis for the medians of the sone intervals is significant at the 0.001 level, while a rank-order analysis of the medians of the lambda intervals is not statistically reliable (P > 0.8). For comparison, Fig. 1 and 2 show *O*s' median judgments for Groups I and II replotted against lambda intervals (the upper curves in Fig. 1) and sone intervals (the lower curves in Fig. 2) respectively. Now the curves for Group I tend to come together while those for Group II are spread apart.

TABLE I

VALUES OF LOUDNESS INTERVALS USED WITH GROUPS I AND II

	Group I		Group II	
Sector	Intervals in sones	Intervals in db. (0.0002 dynes/cm²)	Intervals in lambda units	Intervals in db. (0.0002 dynes/cm²)
Lowest	5	30.0–64.7	0.85	30.0–50.3
	10	35.8–74.4	1.70	35.8–70.1
	15	30.0–79.7	2.55	30.0–77.5
	20	30.0–83.7	3.40	30.0–85.3
Lower middle	5	63.3–73.3	0.85	63.3–74.5
	10	50.0–76.0	1.70	50.0–76.1
	15	40.0–80.1	2.55	40.0–80.7
	20	40.0–84.1	3.40	32.0–85.8
Upper middle	5	79.2–83.3	0.85	75.1–83.3
	10	73.3–83.3	1.70	64.3–83.3
	15	50.0–81.0	2.55	40.6–81.0
	20	58.1–85.7	3.40	34.0–86.2
Highest	5	83.3–86.6	0.85	79.3–86.6
	10	79.2–86.6	1.70	69.9–86.6
	15	73.3–86.6	2.55	56.3–86.6
	20	63.3–86.6	3.40	35.4–86.6

With which scale are *Os'* ratio-judgments linear? Inspection of Fig. 1 and 2 indicates that the ratio-judgments do not conform to either the sone or the lambda scales. Because equal intervals on the sone scale do not

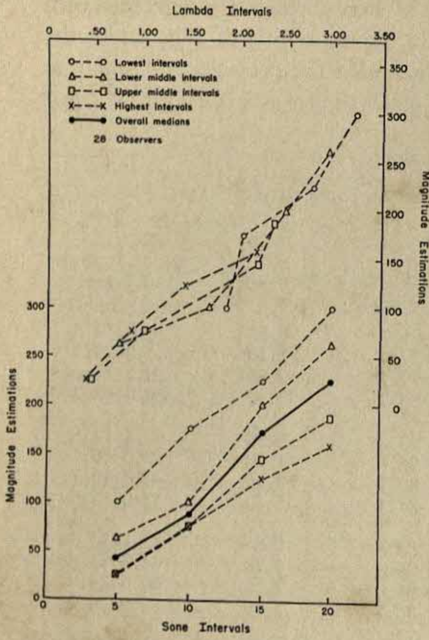


FIG. 1. MEDIAN MAGNITUDE-ESTIMATIONS FOR GROUP I

The medians of *Os'* median estimations of the loudness-intervals in Group I plotted against intervals in sones (bottom abscissa, left ordinate) and intervals in lambda units (top abscissa, right ordinate). The median judgments of the intervals selected from each of the four sectors of the sone scale (See Table I) and the over-all medians are shown by the lower set of curves. The data replotted with respect to lambda intervals are shown by the upper set of curves.

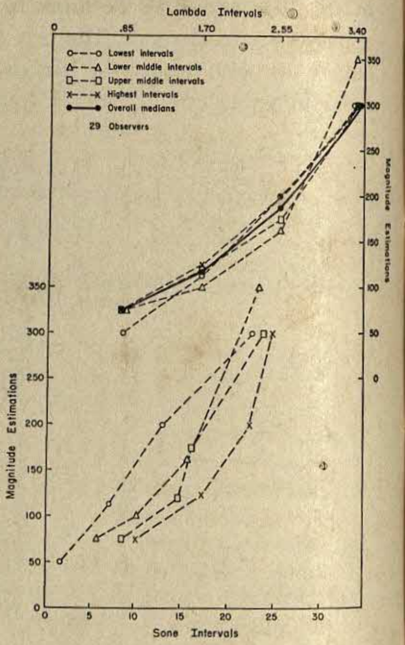


FIG. 2. MEDIAN MAGNITUDE-ESTIMATIONS FOR GROUP II

The medians of *Os'* median estimations of the loudness-intervals in Group II plotted against intervals in sones (bottom abscissa, left ordinate) and intervals in lambda units (top abscissa, right ordinate). The median judgments of the intervals selected from each of the four sectors of the lambda scale (See Table I) and the overall medians are shown by the upper set of curves. The data replotted with respect to sone intervals are shown by the lower set of curves.

yield similar judgments, there is no simple relationship between *Os'* median judgments and the scale values of the intervals on the sone scale. *Os'* judgments plotted against the lambda scale, however, follow an identifiable pattern. Intervals separated by equal distances on the lambda scale become separated by approximately equal ratios, causing the curve to be positively accelerated. For example, the over-all median judgments for Group II are 75, 116.5, 187.5 and 300. Intervals separated by 0.85 lambda units are therefore separated by a ratio of approximately 1.6. For Group I, separations of 0.85, 1.70, and 2.55 lambda units correspond to values of 67, 115, and 210; distances of 0.85 lambda units are here separated by ratios of 1.7 and 1.8 respectively. The agreement of the results for Groups I and II is shown when the data are plotted (in Fig. 3) on a log scale against *jnds* calculated from Riesz's intensity-discrimination data.[5] With the exception of the two smallest *jnd* intervals in Group I, straight lines with similar slopes describe the relationship between the logarithm

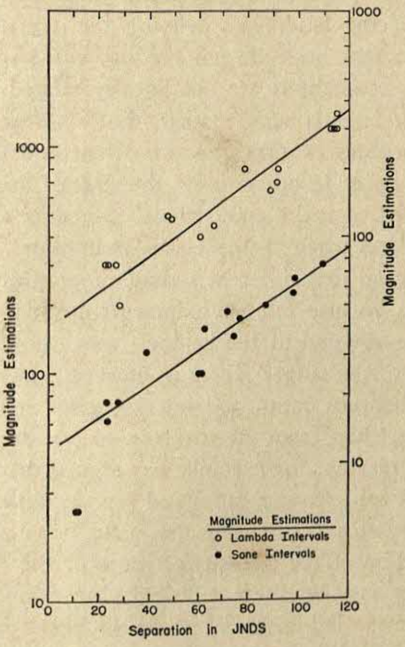


FIG. 3. MEDIAN MAGNITUDE-ESTIMATIONS OF GROUPS I AND II AGAINST SIZE OF INTERVAL IN *Jnds*

The median judgments of the sone intervals are plotted on the left ordinate and of the lambda intervals on the right ordinate.

---

[5] R. R. Riesz, Differential intensity sensitivity of the ear for pure tones, *Phys. Rev.*, 31, 1928, 867–875; The relation between loudness and minimum perceptible increment of intensity, *J. acoust. Soc. Amer.*, 4, 1933, 211–216.

of *O*s' ratio-estimations and the number of *jnd*s contained in the loudness interval. Thus, the data indicate that equal increments of *jnd*s become separated by approximately equal ratios.

Discussion. The results of the experiment show (a) that equal distances on the lambda scale produce roughly equal judgments, regardless of the absolute values of the loudnesses defining the interval; (b) that equal differences on the sone scale do not produce equal judgments; and (c) that the *O*s' ratio-estimations are not linearly related to interval-size on either the sone or lambda scale—rather, that a linear relationship holds between the logarithms of *O*'s ratio-estimations and the sizes of the intervals in lambda units. In other words, the sizes of the loudness-intervals, when *O*s are asked to make equisectional estimations, are a logarithmic function of their sizes when *O*s make ratio-judgments.

Stevens has argued that different scaling procedures lead to conflicting scales of loudness because interval-judgments are biased; differences that are distinct at the low end of the loudness range become much less clear in the upper part of the range.[6] There is, however, evidence to suggest that the discrepancy between ratio- and interval-scales stems from the judgmental task rather than from the tendency to confuse distance with discriminability.[7] Torgerson, for example, has argued that *O*s find it difficult to distinguish between sensory ratios and sensory differences of loudness. Thus, the same sensory relationship that yield loudness-judgments which are a power function of the intensity of sound if the *O*s are instructed to make ratio-judgments of loudness, will yield a logarithmic function if the *O*s are instructed to make judgments of intervals.[8] These two hypotheses make different predictions in the present study. If interval-scales are biased because of an asymmetry in sensitivity, the method of judgment should not affect the results obtained. That is, ratio-judgments of intervals should be linearly related to the sizes of intervals determined by the method of equisection. If, however, the discrepancy between ratio- and interval-scales arises from the task of the *O* (*i.e.* whether the *O* is asked to make a distance judgment or a ratio-judgment) equisection judgments of intervals should turn out to be a logarithmic function of ratio-judgments of intervals. The fact that the size of the intervals in lambda units (determined by the method of equisection) turned out to be a logarithmic function of their ratio-estimations is, therefore, in agreement with Torgerson's conjecture.

---

[6] Stevens, Ratio scales, partition scales, confusion scales, *op. cit.*, 49–66.

[7] W. S. Torgerson, Distances and ratios in psychophysical scaling, *Acta Psychol.*, 19, 1961, 201–205.

[8] *Idem.*

The results suggest that *O*s interpret *jnd*-differences between loudness intervals as distances if instructed to make judgments of equal-appearing intervals, and as ratios if instructed to make judgments of the ratio of loudness-intervals. What is implied is that for the scaling of loudness-intervals, ratio and interval scaling-methods represent alternative ways of describing a single sensory relationship. Beck and Shaw have previously argued that ratio-judgments of pitch should not be interpreted as quantitative measures of sensory ratios but as indicating how tonal relationships, which are in themselves non-quantitative, are mapped onto a number-scale.[9]

SUMMARY

Ratio-estimations of loudness-intervals selected from the sone scale and from the lambda scale were obtained. The results show: (a) that equal differences on the lambda scale produce roughly equal judgments regardless of the absolute values of the loudness defining the interval; (b) that equal distances on the sone scale do not produce equal judgments; and (c) that *O*s' ratio judgments are not linearly related to interval size on either the sone or lambda scale. Rather, a linear relationship holds between the logarithm of *O*s' ratio-estimations and the sizes of the intervals in lambda units. Thus, the results suggest that *O*s interpret the same *jnd*-differences between loudness-intervals as a distance or as a ratio depending on whether they are asked to make equisectional judgments or judgments of the ratios of loudness-intervals. What is implied is that for the scaling of loundness-intervals the methods of ratio- and interval-scaling represent alternative ways of describing a single sensory relationship.

---

[9] Jacob Beck and W. A. Shaw, The scaling of pitch by the method of magnitude-estimation, this JOURNAL, 74, 1961, 242–251; Magnitude estimations of pitch, *J. acoust. Soc. Amer.*, 34, 1962, 92–98; Single estimates of pitch magnitude, *ibid.*, 35, 1963, 1722–1724.

EFFECTS OF HIPPOCAMPAL ABLATION AND INTERTRIAL INTERVAL ON ACQUISITION AND EXTINCTION IN A COMPLEX MAZE

By LEONARD E. JARRARD and THOMAS C. LEWIS, Washington and Lee University

If the hippocampus plays an important role in recent memory, it should be expected that performance of hippocampectomized animals would be affected by length of the intertrial interval. A recent test of this yielded ambiguous results: in acquisition, the spacing of trials had no effect on the performance of the operates; in extinction, the group with the longer intertrial interval showed greater resistance to extinction.[1] The animals were studied in a runway, and the possibility should be considered that this rather simple situation is insufficiently sensitive to the effects of the lesion. In the present experiment, a complex maze was used in which initial learning, retention, relearning and extinction were studied.

METHOD

Subjects. The *Ss* were 72 male rats between the ages of 110 and 130 days: 36 albino rats of the Dublin-Disease-Resistant strain and 36 hooded rats of the Long Evans strain. The animals of each strain were divided into three groups of equal size: an Unoperated Control Group; a Hippocampal Group; and an Operated-Cortical Control Group.

Surgical and histological procedures. The animals were anesthetized with pentobarbital (40 mg/kg), a mid-line scalp incision was made, and suitable holes bored in the skull. For *Ss* of the Hippocampal Group, a glass pipette connected to an aspirator was used to remove cortical tissue overlying the hippocampus and then the exposed hippocampus was removed both medially and ventrally; for those of the Cortical Control Group, cortical ablation was similar to that in the hippocampal *Ss* except a greater amount of cortical tissue was removed in an attempt to equate for total amount of brain damage. The neocortical lesions included damage

---

* Received for publication May 20, 1966. The participation of T. C. Lewis in the research project was made possible by funds from the National Science Foundation Undergraduate Research Participation Program G-140. The authors wish to thank K. A. Marion, A. D. Smith, and R. L. Stump for their assistance with the animals.

[1] L. E. Jarrard, R. L. Isaacson, and W. O. Wickelgren, Effects of hippocampal ablation and intertrial interval on runway acquisition and extinction, *J. comp. physiol. Psychol.*, 57, 1964, 442–444; for studies suggesting a recent memory-function see B. Milner and W. Penfield, The effect of hippocampal lesions on recent memory, *Trans. Amer. Neurol. Assn.*, 80, 1955, 42–48; and L. S. Stepien, J. P. Cordeau, and T. Rasmussen, The effect of temporal lobe and hippocampal lesions on auditory and visual recent memory in monkeys, *Brain*, 83, 1960, 470–489.

to the underlying cortical fiber-layer between cortex and the ventricular surface of the hippocampus.

After the experiment the operates were sacrificed and perfused intracardially with 10% formalin. The brains were photographed and then dehydrated and imbedded in paraffin. Sections were made at 15 micra and every twentieth section through the lesion was mounted and stained with thionin.

Apparatus. The apparatus consisted of an 8-cul Lashley Type-III maze with retrace doors. The four parallel alleys were 38 in. long × 4 in. wide. Microswitches were arranged on the start- and goal-box doors in order to provide measures of latency (time spent in the start box) and total running time. Errors were recorded whenever *S* entered a cul-de-sac or retraced a distance of one-half body length.

Preliminary training. On the fourteenth postoperative day the *S*s were placed on a deprivation-schedule of 10 gm. of ground Purina laboratory chow daily. Preliminary training was started seven days later and consisted of giving each *S* two 45-mg. food pellets in the goal-box for 10 trials on each of three successive days. The second phase of preliminary training consisted of placing *S* in an enclosed runway with food pellets in the food cup at the opposite end. The runway was fitted with a retrace door similar to those used in the maze. Each *S* was given 10 trials a day for two days in the runway.

For both strains, each of the three treatment-groups (hippocampal, cortical control, and unoperated control) was divided into two subgroups, matched for performance in the runway during preliminary training. These subgroups formed the two intertrial interval conditions of 10 sec. and 10 min. Thus, for each strain there were six combinations of treatment and intertrial interval.

Experimental training. The albino animals, in the first stage of the experiment, were given 10 acquisition-trials a day in the Lashley maze at either the 10-sec. or 10-min intertrial interval. The *S*s were run to a criterion of 10 consecutive errorless trials, or for a maximum of 15 days (total of 150 trials). Forty-two days after completion of the first stage of training, retention and relearning were tested by giving the *S*s an additional three days of training under the same conditions that had been used in acquisition. During the time between acquisition and relearning, the *S*s were maintained on an *ad lib.* feeding schedule until the beginning of the deprivation-schedule seven days before relearning. Following relearning, the animals received 10 nonreinforced trials a day under their previous intertrial interval condition to an extinction-criterion of three consecutive trials in which the goal-box was not reached within 60 sec. or until the end of the 5-day extinction-period.

The hooded rats were given five days of acquisition-trials in the first stage of training with the procedure in all other respects being identical to that used with the albino *S*s. Then, starting on Day 6, the running response was extinguished as before. Thus, for the hooded rats, the experimental procedure was analogous to that used in the earlier runway study.[3]

RESULTS

(1) Histological. Examination of the brain sections revealed extensive

---

[3] Jarrard, Isaacson, and Wickelgren, *op. cit.,* 442–444.

bilateral lesions involving from 45 to 85% destruction of the hippocampus in all *S*s in the Hippocampal Group. In the animals with only neocortical lesions, examination of the sections revealed more neocortex destroyed than in animals with hippocampal lesions. The amount of neocortical tissue destroyed in the cortical control *S*s ranged from 10 to 15%. Some thalamic gliosis in the lateral geniculate nucleus was present in several hippocampal and several cortical control *S*s but no correlation was found between the amount of gliosis or size of the lesions and performance of the animals. Tracings of selected sections showing the lesions in a representative hippocampal animal and one with control neocortical lesions are illustrated in Fig. 1.

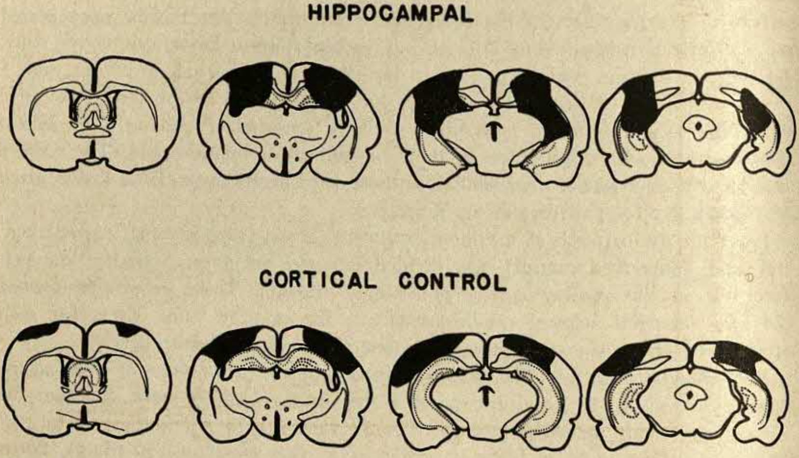


FIG. 1. TRACINGS OF BRAIN SECTIONS SHOWING THE LOCATION AND EXTENT OF LESIONS IN REPRESENTATIVE SUBJECTS WITH HIPPOCAMPAL AND NEOCORTICAL CONTROL DAMAGE

(2) Experimental. The basic data for acquisition consisted of number of trials to criterion (for albino animals), number of errors, latency, and running time. For the albino rats, mean number of trials to criterion for initial learning were 93.5, 58.1, and 51.2 for the Hippocampal, Cortical Control, and Unoperated Control Groups, respectively. A score of 150 was assigned to four hippocampal *S*s that did not reach the criterion. Analysis of variance of these data revealed significant differences among the groups ($F = 14.18$, $df = 2/14$, $p < 0.01$); neither the intertrial-interval effect ($F = 0.01$, $df = 1/14$, $p > 0.10$) nor the interaction of groups with intertrial interval ($F = 0.11$, $df = 2/14$, $p > 0.10$) were significant. Analysis of differences between pairs of means showed that

hippocampal *S*s required significantly more trials to reach the criterion than *S*s in either control group. These results showing impaired learning of rats with hippocampal-lesions agree with those reported by Kaada, Rasmussen, and Kveim in a 14-unit T-maze and Kveim, Setekleiv, and Kaada in the Hebb-Williams maze.[3] In view of the original purpose of the experiment, it is of considerable interest that there were no differences between the massed- and distributed-practice of the hippocampal groups in number of trials to criterion.

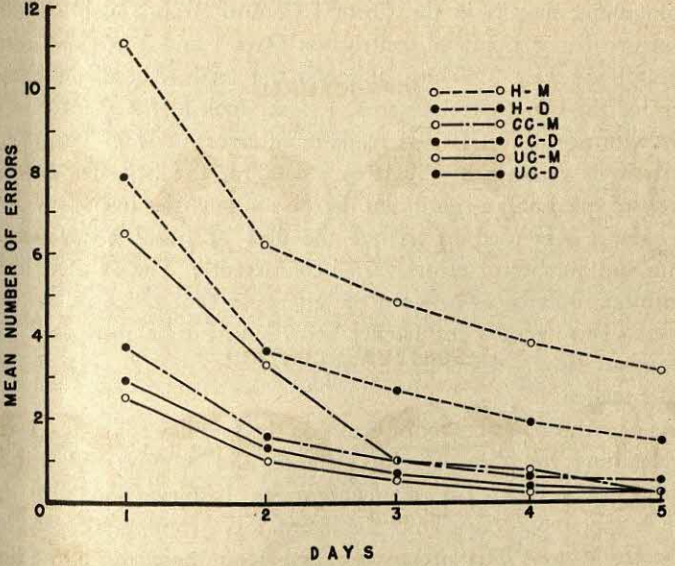


FIG. 2. MEAN NUMBER OF ERRORS ON SUCCESSIVE DAYS OF ACQUISITION

Group designations are H-M = hippocampal-massed; H-D = hippocampal-distributed; CC-M = cortical control-massed; CC-D = cortical control-distributed; UC-M = unoperated control-massed; and UC-D = unoperated control-distributed.

Analysis of error-scores was made by employing number of errors for the first five days of acquisition for both albino and hooded rats. The data were combined, since the procedures were the same and inspection revealed no differences between the two groups of scores. These data are shown graphically in Fig. 2, where mean number of errors have been plotted for successive days. Each point in the figure is based on an N of 12. Analysis of variance of these data showed that the treatment groups differed ($F = 16.59$, $df = 2/66$, $p < 0.01$), there were significant changes over days ($F = 17.01$, $df = 4/12$, $p < 0.01$), and there was a significant interaction between treatment-groups and days ($F = 9.09$, $df = 8/264$, $p < 0.01$). Inspection of Fig. 2 and analysis of differences between means showed, in addition to the obvious differences between hippocampal and control groups, that *S*s in the Cortical Control Group made significantly more errors than unoperated controls on Days 1 and 2 but did not differ significantly by Day 3. Thus, although the trials-to-criterion data were similar for the two control groups, a closer look at the error-data shows that *S*s with neocortical lesions made more errors early in learning.

---

[3] B. R. Kaada, E. W. Rasmussen, and O. Kveim, Effects of hippocampal lesions on maze learning and retention in rats, *Exp. Neurol.,* 3, 1961, 333–355; O. Kveim, J. Setekleiv, and B. R. Kaada, Differential effects of hippocampal lesions on maze and passive avoidance learning in rats, *Exp. Neurol.,* 9, 1964, 59–72.

Analysis of variance of the latency data for the first five days of acquisition revealed only a significant decrease in latencies over days. Analysis of covariance was used to analyze the data of running-time since this measure and number of errors varied concurrently. The analysis indicated no significant differences between the groups in time taken to run through the maze. These results for latency and running time show that *S*s with hippocampal lesions responded as vigorously as control *S*s to the experimental situation.

Tests of retention (number of errors on the first five trials of relearning) and relearning (number of errors, latency, and running time) indicated no differences between the treatment-groups. Lashley found that damage to the hippocampus caused no significant loss of preoperatively learned maze-performance.[4] The present results extend these earlier findings by showing that even though animals with hippocampal lesions experience more difficulty in postoperative learning of a complex maze, they retain the information as well as control animals once it is learned.

A careful inspection of the extinction-data for both albino and hooded *S*s indicated that the data were similar even though the training procedures had differed. Therefore, the extinction-data were combined for purposes of statistical analysis. Means of number of trials to the criterion of extinction were 42.0, 22.0, and 17.6 for the Hippocampal, Cortical Control, and Unoperated Control Groups. Analysis of variance of these data disclosed significant group differences ($F = 55.73$, $df = 2/22$, $p < 0.01$) but no significant effect for the interaction of groups by intertrial-interval. These results agree with those found in other studies in showing that hippocampectomized animals have trouble extinguishing learned responses.[5] The earlier finding, however, of a significant difference between massed- and distributed-practice hippocampal animals in extinction in the runway was not found in the more complex maze-situations. Analysis of the error, latency, and running time for the first day of extinction (before any *S* had extinguished) showed that hippocampal and cortical control groups were significantly faster than unoperated controls in latency and running time but there were no differences in errors.

---

[4] K. S. Lashley, Studies of cerebral function in learning: XII. Loss of the maze habit after occipital lesions in blind rats, *J. comp. Neurol.*, 79, 1943, 421–462.

Discussion. The present results do not support the view that the hippocampus is involved in recent memory-functions, since hippocampal *S*s with distributed practice trails did not differ from hippocampals that received massed trials. It therefore appears that the earlier failure to find differences in runway acquisition in hippocampals with different intertrial intervals cannot be attributed to the relatively simple nature of the task.[6] The finding of similar extinction for massed- and distributed-practice hippocampal *S*s in the present situation does not agree with the results showing slower extinction of hippocampals with distributed-practice in the runway. It is obvious that the two experimental situations differ in a number of ways. Experiments dealing with this problem are currently being undertaken.

It is not clear at the present time whether the impaired learning in animals with hippocampal lesions is due to changes in motivational functions that are closely related to the learning process or to neural mechanisms more immediately involved in learning. Results of several recent studies suggest that rats with lesions in the hippocampus differ from control *S*s in motivation. For example, hippocampectomized rats have been found to be more active in relatively novel situations.[7] It is also possible that hippocampal animals differ from controls in terms of food motivation. Rats with hippocampal lesions do not eat more food than controls but they do spend more time eating in a novel situation when food is available.[8] It seems that changes in motivation following hippocampal ablation must be ruled out as possible explanations before the impaired learning can be attributed to deficits in learning ability or in memory.

---

[5] R. L. Isaacson and W. O. Wickelgren, Hippocampal ablation and passive avoidance, *Science,* 138, 1962, 1104–1106; Jarrard, Isaacson, and Wickelgren, *op. cit.,* 442–444.

[6] Jarrard, Isaacson, and Wickelgren, *op. cit.,* 444.

[7] W. W. Roberts, W. N. Dember, and Malcolm Brodwick, Alternation and exploration in rats with hippocampal lesions, *J. comp. physiol. Psychol.,* 55, 1962, 695–700; Herman Teitelbaum, and Peter Milner, Activity changes following partial hippocampal lesions in rats, *J. comp. physiol. Psychol.,* 56, 1963, 284–289.

[8] L. E. Jarrard, Hippocampal ablation and operant behavior in the rat, *Psychon. Sci.,* 2, 1965, 115–116; the data showing hippocampals spend more time eating in a novel situation are from an unpublished study by the senior author.

SUMMARY

The hypothesis that the hippocampus is involved in recent memory-functions was tested by employing different intertrial intervals in maze acquisition, retention, and extinction. Half of the *S*s in each group (hippocampal operates, cortical controls, and unoperated controls) received trials in a Lashley Type-III maze with a 10-sec. intertrial interval, while the other half had 10-min. intertrial intervals. *S*s with hippocampal lesions were significantly impaired in acquisition and extinction but there was no differential effect with the different intertrial intervals. Retention and relearning of the postoperatively acquired responses after 42 days showed that hippocampal *S*s retained the information as well as controls once it was learned. The present results do not support the view that the hippocampus is involved in recent memory-functions since learning was similarly impaired in hippocampal *S*s that received massed and those that received distributed practice trials. The nature of the impaired learning in animals with hippocampal lesions is not known at the present time.

EFFECTS OF VARIATION IN SHOCK-INTENSITY ON THE BEHAVIOR OF DORSAL-HIPPOCAMPECTOMIZED RATS IN TWO PASSIVE-AVOIDANCE SITUATIONS

By JOHN J. BOITANO, Rosary Hill College, and ROBERT L. ISAACSON, University of Michigan

Kaada, Rasmussen, and Kveim have shown that rats with hippocampal lesions learn as rapidly as normals to withhold an approach-response to an electrified water-dish,[1] but Isaacson and associates have shown that other rats with these lesions are deficient in another kind of passive-avoidance learning.[2] Kaada made lesions by D.C.-fulguration of the anterior and posterior hippocampus, while Isaacson produced radical hippocampal ablation by aspiration. Kaada's testing situation permitted free responding with self-administered mouth-shocks of low intensity (0.18 ma. A.C. in water-deprived *Ss*, while Isaacson studied the effect on food-rewarded discrete-trials performance of a single 15-sec. foot-shock of high intensity (0.52 ma. A.C.). The present experiment was designed to evaluate the importance of some of the procedural differences. Small dorsal hippocampal lesions were made by D.C. fulguration, and their effects were examined in tests involving both high and low intensities of shock.

METHOD

Subjects. The *Ss* were 31 male albino rats, 90–150 days old, and weighing an average 295 gm. at the time of surgery. In 14 *Ss*, there was electrolytic bilateral destruction of the dorsal hippocampus. Anodal current (1.5 ma. D.C. for 45 sec.) was passed through an electrode positioned 2 mm. posterior to bregma, 2 mm. lateral from the midline, and 3 mm. below the dura. The cathodal electrode was positioned in the anus. The other 17 animals served as controls. They were anesthetized and placed in a Johnson stereotaxic instrument, the electrode was lowered to the dorsal hippocampus, but no current was delivered. The random assignment of all animals to two levels of shock-intensity (0.18 and 0.51 ma.) resulted in four groups of *Ss*: a low-shock control group ($N = 9$); a low-shock group with

---

* Received for publication June 12, 1966. This research was completed while the senior author was a participant in a National Science Foundation Research Participation Program at the University of Michigan.

[1] B. R. Kaada, E. W. Rasmussen, and Olav Kveim, Impaired acquisition of passive avoidance behavior by subcallosal, septal, hypothalamic, and insular lesions in rats, *J. comp. physiol. Psychol.*, 55, 1962, 661–670.

[2] R. L. Isaacson and W. O. Wickelgren, Hippocampal ablation and passive avoidance, *Science*, 138, 1962, 1104–1106; D. R. Snyder and R. L. Isaacson, Effects of large and small bilateral hippocampal lesions on two types of passive-avoidance responses, *Psychol. Rep.*, 16, 1965, 1277–1290.

hippocampal lesions ($N = 7$); a high-shock control group ($N = 8$); and a high-shock group with hippocampal lesions ($N = 7$). All the animals were allowed 12–14 days for recovery before being placed on water-deprivation.

Histology. After the experiment, the *Ss* were perfused intracardially with 0.9% saline followed by a 10% solution of formalin. After dehydration and embedding in paraffin, the brains were sectioned coronally at 40 μ. Histological examination of the thionin-stained brains revealed (in the majority of animals) bilateral destruction of the dorsal hippocampus, involving its rostral and caudal portions, just anterior to the ventral and lateral curvature (see Fig. 1). The amount of bilateral dorsal destruction ranged from 49–88% with a median value of 70%. Three animals showed unilateral destruction of the dorsal hippocampus, the destroyed side being damaged in excess of 61% while the undestroyed side was damaged, on the average, only 6%. These animals were not significantly different from the bilateral ablates in any of the behavioral tasks employed. Cortical damage usually was restricted to an area adjacent to the midline which left the banks of the cingulate gyrus intact. Typically, this area extended to 2–3 mm. behind bregma and included an area from 1–3 mm. laterally. In only one animal did the damage involve bilateral destruction of the middle cingulate cortex. No direct thalamic involvement was found, although several instances of gliosis of the lateral thalamic nuclei were noted.

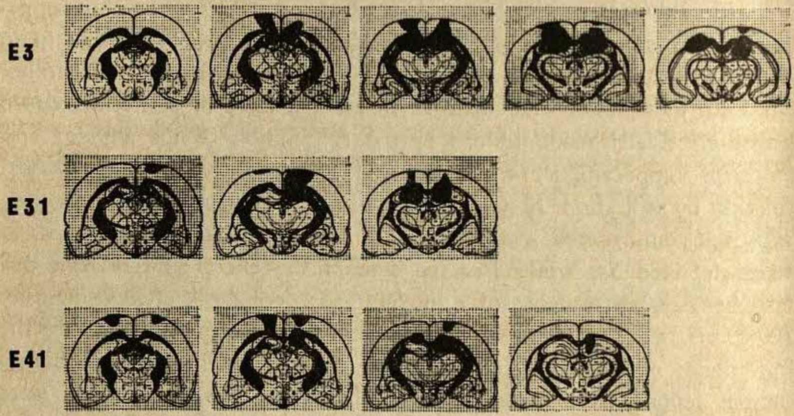


FIG. 1. RECONSTRUCTION OF LESIONS IN THREE REPRESENTATIVE ANIMALS (*E*3, an example of a good bilateral lesion; *E*31, an intermediate ablation; *E*41, a predominantly unilateral lesion.)

Apparatus. For the continuous avoidance (*CA*) task, the apparatus was similar to that of Snyder and Isaacson.[3] It consisted of a single plywood chamber (22 × 20 × 10 in.) with a floor of copper. A section of heavy cage-paper was placed over the top as a ceiling. In the front of the box, a red-cellophane-covered Plexiglas door (9 × 6 in.) permitted undetected visual observation of each *S* during testing. A house-light (25 w.) mounted on the back wall was the sole source of illumination. An electrified water-dish was placed in the front-right corner, adjacent to the door, and a matched-impedance shock-source was employed to deliver shock whenever *S* closed the circuit by licking the electrified water. Each circuit-closure was registered by an event recorder. An ultrasonic activity-device was mounted on the center of the side wall opposite the water-dish and 4 in. below the ceiling.[4] Through one of the attached transducers, an ultrasonic signal (40 K cps) was radiated into the confines of the experimental chamber. As the animal disrupted the nodes of standing waves by moving, a second transducer detected the regional change which caused a glow-tube to fire and a counter to be activated. The sensitivity of the instrument was such that only rapid gross movements were recorded; slow or very circumscribed movements, like preening or deliberate trunk-extensions, were not sufficient to fire the counter, but walking, running, rapid stretching, sniffing, rapid preening, and turning were recorded.

---

[3] Snyder and Isaacson, *op. cit.*, 1280.

For the discrete-trials avoidance (*DTA*) procedure, the apparatus was the same as that used by Isaacson and Wicklegren.[5] It consisted of a two-chamber box with the larger chamber (12 × 18 × 10 in.) separated from the smaller chamber (6 × 8 × 4 in.) by a guillotine door. A small drinking cup was permanently mounted to the wooden floor of the larger chamber directly opposite the entrance-way to the smaller chamber. The smaller compartment had a copper-grid floor which terminated beneath a drinking cup on its far wall. Thus, the two drinking cups were located at the extreme ends of the apparatus in a direct line with the entrance-way to the small compartment. The roof of the small chamber was ¼ in. hardware cloth, while the ceiling of the larger compartment consisted of clear Plexiglas. A 100-w. lamp 10 in. above the small chamber provided the sole source of illumination during testing.

Procedure. Half of the animals in each group were tested first in the *DTA*-task, while the remaining animals were tested first in the *CA* task. In the *DTA* condition, each *S* was trained under 23 hr. of water-deprivation to enter the smaller compartment within 30 sec. after the door was raised in order to obtain a water-reward (4 drops). Then it was lured back into the larger compartment by a smaller reward (one drop), and the guillotine door lowered. Each such cycle constituted a trial. A 1-min. intertrial interval was used. There were 10 trials per day for six consecutive days. On the 35th trial, shock was administered for 15 sec. through the grid floor of the smaller chamber while the animal was drinking. The entrance-latencies into the smaller chamber were measured on all trials.

In the *CA* condition, each *S* was placed initially in the experimental chamber for 3 min. of free exploration under 48 hr. of water-deprivation. At the end of this time, *S* was removed, the electrified water dish was inserted, and *S* was replaced. Testing in the presence of the electrified water-dish lasted 20 min., with activity-scores recorded after each 5-min. interval. The number of self-administered mouth-shocks was automatically recorded, and two types of approach-withdrawal behavior were observed through the Plexiglas door. The first was that defined by Naess and Rasmussen—a slow, cautious approach to the immediate vicinity of the dish with a rapid backward shuffling of the hind legs away from the dish.[6] The orientation of the head and trunk was towards the dish in both the approach and the withdrawal aspects. The second type of behavior was a sidewise approach with an elongated trunk-extension bringing *S* up to but not quite in contact with the dish. The withdrawal-phase consisted of a simultaneous retreat and readjustment of the bodily orientation towards the side wall, followed by a slight pause, and then a full gallop towards the rear of the test chamber. If the pause was longer than momentary (1 sec.), the response was not counted. In both types of responses, the hind legs had to be in motion during the withdrawal-phase.

---

[4] L. J. Peacock and Marshall Williams, An ultrasonic device for recording activity, this JOURNAL, 75, 1962, 648–652.

[5] Isaacson and Wicklegren, *op. cit.,* 1104.

RESULTS

Table I presents the median number of mouth-shocks received by each group and the total number of approach-withdrawal responses in the *CA*-situation. Although both hippocampal groups received fewer shocks than did the controls, the variance due to groups was not statistically significant either with ($F < 1$) or without ($F = 2.09$, $df = 1/27$, $p > 0.05$) a log $(1 + x)$ transform. The variance due to shock-level also was significant (without the transform, $F = 3.03$, $df = 1/27$, $p > 0.05$; with it, $F = 1.77$), as was the interaction ($F = 1.48$; $F < 1$). Analysis of the number of approach-withdrawals, with and without transformation, yielded no significant effects (all *F*-values < 1).

TABLE I

SHOCKS RECEIVED AND APPROACH-WITHDRAWALS IN *CA*

		No. of shocks		No. of approach-withdrawals	
Group	*N*	Mdn.	Range	Mdn.	Range
Control—Low Shock	9	25.0	6–460	13.0	2–32
Hippo. —Low Shock	7	15.0	8–22	12.0	5–24
Control—High Shock	8	7.5	2–13	13.0	3–22
Hippo. —High Shock	7	5.0	1–12	15.8	0–26

Exposing each animal to the electrified water dish resulted in a post-shock decline in overt activity as indicated in Fig. 2. While the high-shock groups showed less activity than the low-shock groups, the difference was insignificant. No significant differences were found between the groups of lesioned animals.

The data summarized in Fig. 3 indicate that the entrance-latencies of all animals in all groups declined over the course of the first 35 training in the *DTA*. Following the administration of shock, the time required to reënter the smaller compartment increased dramatically for all groups (trials 36–40), and then slowly diminished over trials 41–60. A trend-analysis performed on the reciprocals of the means of blocks 8–12 for all animals in all groups revealed statistically significant differences for the three main order-effects: type of lesion ($F = 6.56$, $df = 1/27$, $p < 0.05$); shock-intensity ($F = 8.65$, $df = 1/27$, $p < 0.01$); trials ($F = 12.74$, $df = 4/108$, $p < 0.01$) as well as the shock-by-trials interaction ($F = 2.46$, $df = 4/108$, $p < 0.05$).[7] Animals with hippocampal destruction had longer entrance latencies under both shock levels than did their respective control groups.

---

[6] K. Naess and E. W. Rasmussen, Approach-withdrawal responses and other specific behavior reactions as screening tests for tranquilizers, *Acta Pharmacol. Toxicol., Kbh.*, 15, 1958, 99–114.

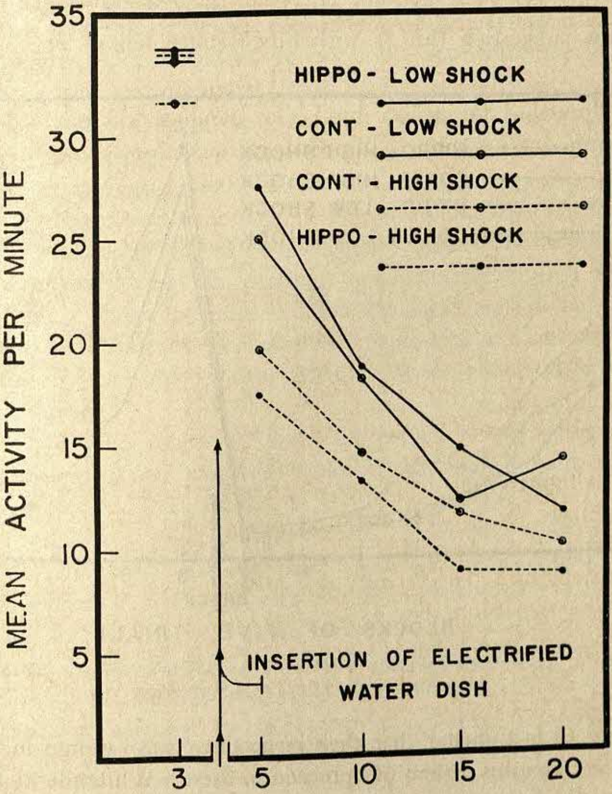


FIG. 2. EFFECTS OF MOUTH-SHOCK ON ACTIVITY IN THE *CA*-TASK

---

[7] A. L. Edwards, *Experimental Design in Psychological Research,* rev. ed., 1960, 233–239.

Discussion. The results suggest that rats with electrolytic lesions of the dorsal hippocampus are not deficient in the passive avoidance of shock, whether of high or low intensity. By contrast, rats with massive hippocampal destruction have been found to reënter readily a compartment in which the shock had been received.[8] Snyder and Isaacson reported that destruction of the dorsal hippocampus by aspiration produced some impairment in a *DTA*-situation but very little effect in *CA*.[9] There is, moreover, some suggestion that *S*s with hippocampal lesions of the present study were, in fact, 'better' than their respective control groups in avoiding the noxious stimulus. In the *CA* procedure, there was a tendency for both hippocampal groups to accept fewer mouth shocks and exhibit less activity in the presence of the electrified water-dish than their controls. Similar results were obtained in the *DTA*-task. The hippocampal *S*s avoided the smaller compartment for significantly longer periods than the controls.

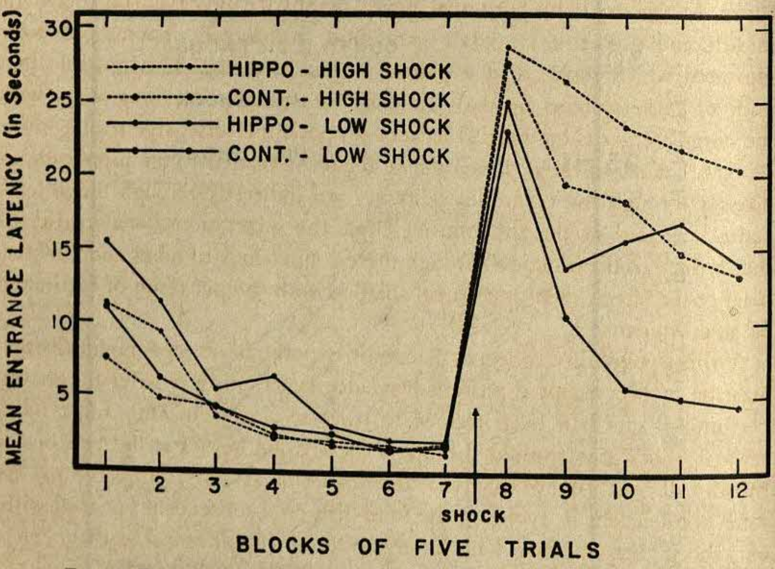


FIG. 3. ENTRANCE-LATENCY (IN SEC.) AS A FUNCTION OF TRIALS AND SHOCK IN THE *DTA*-SITUATION

There are at least three possible explanations of the divergence of these results from those obtained in previous studies. The lesions in the present study were created by *D.C.*-fulguration rather than by aspiration. There has been a recent suggestion that certain behaviors resulting from neural tissue destruction may in effect be an artifact of the lesion-making process.[10] For instance, in the electrolytic process, the introduction of metallic ions generally produces scar tissue which may have chronic irritative effects on the surrounding tissue. Thus, the source of any observed behavioral change may be attributable either to tissue-destruction, or to the formation of irritative scar tissue, or both.

---

[8] Isaacson and Wicklegren, *op. cit.,* 1105.

[9] Snyder and Isaacson, *op. cit.,* 1281, 1288.

A second explanation involves the bilateral sparing of the cingulate cortex concomitant with the dorsal hippocampal destruction of the present study. In the small hippocampal group of Snyder and Isaacson, both the middle and posterior cingulate gyrus and the dorsal hippocampus were aspirated which resulted in a passive-avoidance deficit. It is possible that both of these systems normally interact in mutually opposing ways with the cingulate gyrus tending to balance hippocampal discharge to the lower centers.[11] Anatomically, the cingulate region is related to the hippocampus directly through the cingulum posterius and indirectly through the anteroventral nucleus of the thalamus.[12] Thus, the intact transitional cortex of the midline could function through these connections to offset and counterbalance the effects of hippocampal ablation with the net result of facilitated passive avoidance.

A third possibility concerns Kimura's hypothesis of topographical organization of hippocampal passive-avoidance functioning.[13] After hippocampectomized rats had been trained to traverse a straight alley for a food-reward, shock was applied through a metal food-holder at the end of the runway. It was found that the anterior-lesioned dorsal-hippocampal *Ss* were not deficient in avoiding the food-dish on a subsequent test-trial without shock, but animals with more posterior hippocampal damage were deficient and readily traversed the runway. With slight overlap, Kimura's anterior group corresponds to those of the lesions in the present study, while the damage sustained by his posterior group appears to have involved the lateral and ventral curvature of the hippocampus. These results in conjunction with our own suggest that some parts of the hippocampus may be more responsible than others for the release and control of fear-motivated behavior.

---

[10] R. W. Reynolds, An initative hypothesis concerning the hypothalamic regulation of food intake, *Psychol. Rev.*, 72, 1965, 105–116.

[11] L. K. Gerbrandt, Neural systems of response release and control, *Psychol. Bull.*, 64, 1965, 113–123; E. C. Crosby, Tryphena Humphrey, and E. W. Laurer, *Correlative Anatomy of the Nervous System*, 1962, 475.

[12] J. W. Papez, A proposed mechanism of emotion, *Arch. Neurol, Psychiat.*, 38, 1937, 725–743; W. J. H. Nauta, An experimental study of the fornix system in the rat, *J. comp. Neurol.*, 104, 1956, 247–271; J. E. Rose and C. N. Woolsey, Structures and relations of limbic cortex and anterior thalamic nuclei in rabbit and cat. *J. comp. Neurol.*, 19, 1948, 279–348.

[13] Douglas Kimura, Effects of selective hippocampal damage on avoidance behavior in the rat, *Canad. J. Psychol.*, 12, 1958, 213–218.

SUMMARY

Rats with electrolytically produced bilateral lesions of the dorsal hippocampus and sham-operated controls were exposed to two types of passive-avoidance situation. The animals were not deficient in either situation. There was some indication, in fact, that the hippocampal operates were 'better' in avoiding shock than were the controls.

CONVERGENT ASSOCIATIONS IN MEDIATED GENERALIZATION

By DANNY R. MOATES and JAMES H. KOPLIN, Vanderbilt University

---

The mediation-hypothesis states that the gap between observable inputs and outputs is bridged by an implicit response having stimulus-consequences. The task currently facing mediation-theorists is an explication and manipulation of the implicit response in well-defined experimental situations. Using a mediated generalizaton-task, the following experiment investigates the influence of a single variable, the number of times which a specific implicit response is elicited.

Mink has introduced a reliable design for mediated generalization in which a motor response served as an index of mediated facilitation.[1] Mink used a response equivalence mediation paradigm.[2] In the first stage, the subjects (*Ss*) are shown a series of words and instructed to press a lever upon the appearance of each word. The mediation-hypothesis asserts that the presentation of each of these words implicitly arouses an associated word. The pressing of the lever is thus paired with this implicit associate as well as with the stimulus-word. Following the first list, a second list of words (test-stage) is presented, composed of all words in the first list plus a strong associate to each of these words and a set of control words. The words are presented serially and *S* instructed to press the lever each time a word from the first list appears. Presentation of the strongly associated words typically elicits the lever-press with a frequency significantly greater than that of the control words. Mink's work has been confirmed by other investigators.[3]

The next important step is the identification of the variables which affect the amount of generalization observed. It should be possible to facilitate elicitation of the implicit response by increasing the number of associated stimulus-words, as inferred from word associative norms. For example, the stimulus-words DARK, HEAVY, and MATCH all converge on the associate LIGHT and would facilitate the false recognition of LIGHT more than if the single stimulus DARK were used. Although this hypothesis has not been tested in mediated generalization, Underwood has demonstrated the effect in the recognition of words.[4]

---

* Received for publication August 12, 1966.

[1] W. D. Mink, Semantic generalization as related to word association, *Psychological Reports,* 12, 1963, 59–67.

[2] J. J. Jenkins describes eight three-stage paradigms for investigating mediation effects of which the response equivalence paradigm is one. (J. J. Jenkins, Mediated associations: paradigms and situations, in C. N. Cofer and B. S. Musgrave (eds.), *Verbal Behavior and Learning: Problems and Processes,* 1963, 216–217.)

[3] S. A. Moss, A study of the semantic generalization effect, Unpublished Doctoral dissertation, University of California, 1960; Jenkins, *op. cit.,* 215–216.

[4] B. J. Underwood, False recognition produced by implicit verbal responses, *J. exp. Psychol.,* 70, 1965, 122–129.

The following experiment was designed to extend Underwood's results by examining the effect of varying the number of convergent associations (zero, one, two, or three) presented in the first stage on the degree of mediated generalization observed in the test-stage. The hypothesis was that the more frequently an implicit response was elicited in the first stage, the greater would be the generalization to that word in the test-stage. The number of lever-presses to an associate in the test-list was predicted to be a monotonically increasing function of the number of implicit elicitations of that associate in the first stage.

Method: (1) Subjects. Forty undergraduates enrolled in psychology courses in Vanderbilt Summer Session served as *Ss*. Participation in experiments was a requirement of the course.

(2) Materials. To control for the specificity of single words, four different lists were prepared. Words for the first or learning stage lists (L-words) were drawn from a pool of 12 three-word sets (labeled I through XII), the three words in each set having a common, high-strength associate (A-word). For example, the set RIVER, OCEAN, and BATH all have WATER as a strong associate. The stimulus-words and their associates were drawn from Palermo and Jenkins' associative norms,[5] and from two sets of unpublished norms.[6]

The first list was composed of all 3 words from each of Sets I, II, and III (9 words altogether), 2 words randomly selected from each of Sets IV, V, and VI (6 words), 1 word randomly selected from each of Sets VII, VIII, and IX (3 words) and no words from each of Sets X, XI, and XII. In addition to these 18 stimulus-words, the list contained 3 control words (X-words), bringing the total number of items to 21. The X-words were included to mask the purpose of the experiment and to give a measure of the degree of learning in the first stage uncontaminated by the experimental manipulation. Insofar as could be determined from available norms, the control words were not associates to any L-word or to any A-word.

The remaining three lists were so constructed that each of the 12 three-word sets occurred in each of the four conditions (three, two, one, or zero) in the experiment.

Each first-stage list had a corresponding test-list which contained the 21 words in the learning list plus all 12 A-words and a set of 3 control words (Y-words). The Y-words bore no associative relationship to any word in the test-list. The three A-words in each test-list for which there were no L-words in the learning-list (A-zero words) were effectively the same as Y-words. They appeared neither explicitly nor implicitly in the learning list and so were novel in the test-list. Despite this apparent equality, Y-words were introduced to control for any unexpected effect which might be introduced simply by the A-words' being strong associates of several words. Since both A-words and Y-words were the test for generalization, they were controlled for frequency. All such words were listed as AA in the

---

[5] D. S. Palermo and J. J. Jenkins, *Word Association Norms,* 1964, 1–295.

[6] A set of Vanderbilt University free association norms was collected from 250 *Ss* on 240 nouns in 1964. James Greeno has very kindly provided us normative data collected at the University of Minnesota.

Thorndike-Lorge word-count except METAL and GUN which were A.[7] A practice list composed of eight numbers was used to familiarize the *S*s with the task of generalization. Numbers were used in order to minimize interference with subsequent learning.

(3) Procedure. The *S*s were randomly divided into four groups of 10 each, one for each pair of learning and test-lists. Each *S* was seated before a Lafayette memory-drum and shown a two-button panel at his right hand. *S* was told to hold the lower button down during the first phase except at the appearance of each new word on the memory-drum. Then he was to press the upper button quickly and briefly. He was further instructed to try to remember the words for a later stage of the experiment. The first-stage list was presented at a rate of one word every 2 sec. The complete list was presented twice without pause, using two randomizations. The experimenter (*E*) was screened from view by a panel. A pair of lights behind the panel indicated which button *S* was pressing.

At the end of the second randomization, *E* put the test-list on the memory-drum and instructed *S* to do just as he had done before except to press the upper button only when a word appeared which he thought had been a member of the first list.

Four randomizations of the test-list were presented without interruption except for a pause of a few seconds at the end of the second randomization to change the windows of the memory-drum. The words were presented for a period of 1 sec. followed by an intertrial interval of 1 sec. during which the word was covered by the shutter of the memory-drum. *E* recorded each word for which *S* pressed the upper button.

TABLE I

MEAN NUMBER OF RESPONSES IN EACH CATEGORY ON THE FIRST TEST-TRIAL

Word category	List 1	List 2	List 3	List 4	Grand Mean
L—words	15.60	15.50	15.30	16.50	15.72
X—words	2.50	2.10	2.10	2.80	2.38
Y—words	0.20	0.00	0.00	0.10	0.08
A—zero words	0.00	0.20	0.40	0.20	0.20
A—one words	0.20	0.60	0.10	0.60	0.38
A—two words	0.50	0.40	1.00	0.10	0.50
A—three words	0.70	0.70	1.40	1.10	0.98

Results. In the test-procedure four randomizations of the test-list were presented to each *S* to control for any sequence-effects which might appear in a single randomization. It is possible, however, that successive presentations of the test-list interfered with *S*'s memory of the learning list, creating a positive bias in his response to A-words. The data were therefore analyzed first for trial one and then for the combined four trials.

The variable of interest was the number of presses elicited by A-words in the test-list. Table I shows the mean number of responses for each category on the first test-trial. The maximal value possible for each entry is 3.00, except that for L-words, which is 18.00. X-words, which served as a measure of performance in the first stage, have a mean of 2.38. For comparison with other experiments this value can be translated to 79%, which falls within the range of 60–85% reported by Mink and Martin.[8] L-words were responded to 87% of the time, which is slightly higher yet within the same range as the response to X-words.

---

[7] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

An analysis of variance was performed having as a between-*S* source of variance the four learning-lists, and as a within-*S* source the numbers of stimuli eliciting an A-word. L-words, X-words, and Y-words were not included. The within-*S* factor was tested for the assumptions of equality and symmetry of covariance-matrices.[9] Neither assumption was met, X^2 $(df = 30) = 67.05$, $p < 0.001$; X^2 $(df = 30) = 31.39$, $p < 0.001$. The degrees of freedom for Number of Stimuli and its interaction were accordingly reduced through a procedure described by Greenhouse and Geisser.[10] The revised summary table showed a significant effect for Number of Stimuli, $F\ (3, 92) = 12.01$, $p < 0.001$, and for the Lists × Number of Stimuli interaction, $F\ (8, 92) = 2.31$, $p < 0.05$.

It was hypothesized that generalization to the A-words would increase monotonically as the number of L-words for an A-word increased from zero to three. Since the main effect for Number of Stimuli was significant on the first test-trial, a trend analysis was made of this factor. When data from all four lists were combined, the mean number of responses for each level was: A-zero = 0.20; A-one = 0.38; A-two = 0.50; A-three = 0.98. The trend analysis of these means revealed a significant linear component, $F\ (1, 92) = 32.71$, $p < 0.001$. Although the linear trend was quite marked when lists were combined, a graphic analysis indicated that it was not present for each individual list. These differences among lists were the basis for the significant interaction.

Table II shows the mean number of responses for each category on the combined four test-trials. An analysis of variance was performed having as between-*S* sources of variance the four different learning-lists, and as within-*S* sources the number of stimuli eliciting an A-word, and the four randomizations of each test-list. L-words, X-words, and Y-words were not included. Although no differences among lists were predicted, this factor showed marginal significance, $F\ (3, 36) = 2.75$, $p = 0.057$. This effect is attributed to the wide range of associative strengths among L-words and A-words. Though a narrower range was desired, the shortage of associative norms precluded it. The two within-*S* factors were tested for the assumptions of equality and symmetry of covariance matrices. The assumptions were met for Randomizations, but the test for equality failed for Number of Stimuli, $X^2\ (df = 30) = 58.65\ p < 0.005$. The degrees of freedom for Number of Stimuli and its interactions were accordingly reduced. The revised summary table showed three significant sources among the within-*S* factors: Number of Stimuli, $F\ (1, 36) = 11.67$, $p < 0.001$; Randomizations, $F\ (3, 108) = 2.91$, $p < 0.05$; and the Lists × Number of Stimuli interaction, $F\ (3, 36) = 4.07$, $p < 0.05$. Since the main effect for Number of Stimuli was significant for all four trials as well as for the first, a trend analysis was again made of this factor. The combined data from all four lists showed the mean number of responses for each level to be: A-zero $= 0.34$; A-one $= 0.42$; A-two $= 0.64$; A-three $= 0.92$. The trend analysis of these means also revealed a significant linear component, $F\ (1, 36) = 30.23$, $p < 0.001$.

---

[8] Mink, *op. cit.*, 33; J. G. Martin, Mediated transfer in two verbal learning paradigms, Unpublished Doctoral dissertation, University of Minnesota, 1960.

[9] B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 369–374.

[10] S. W. Greenhouse and Seymour Geisser, On methods in the analysis of profile data, *Psychometrika*, 24, 1959, 95–112.

TABLE II

MEAN NUMBER OF RESPONSES IN EACH CATEGORY ON ALL FOUR TEST-TRIALS

Word category	List 1	List 2	List 3	List 4	Grand Mean
L-words	15.15	14.40	13.85	16.08	14.87
X-words	2.18	1.95	1.82	2.68	2.16
Y-words	0.12	0.05	0.15	0.15	0.12
A-zero words	0.05	0.22	0.88	0.22	0.34
A-one words	0.38	0.58	0.15	0.58	0.42
A-two words	0.58	0.65	1.02	0.30	0.64
A-three words	0.55	0.58	1.40	1.15	0.92

Discussion. The hypothesis that mediated generalization can be facilitated by increasing the number of elicitations of the implicit response is strongly supported by the highly significant effects found for Number of Stimuli in the two analyses of variance. Moreover, the significant linear components in the two analyses of trend affirm the predicted monotonic relationship. These results suggest that the implicit response is a useful construct which can be manipulated with orderly and predictable results. Further, the articulation of this construct adds credibility to the more general mediation-hypothesis.

The results are not unclouded. Two factors, Lists and Randomizations, were control factors for which no experimental effect was expected. The fact that both showed effects indicates that there are other variables active but not anticipated in this task. L-words in the four lists varied considerably in their associative strength with the A-words. The Lists × Number of Stimuli interaction as well as the marginally significant main effect for Lists is attributed to this variability.

A further complication appears among the Y-words. These words were included in the test-list as controls, to serve as a basis for comparison of generalization to A-words. Since none of the Y-words had appeared in the learning-list, it was felt that the number of responses made to them would represent the base level of false identification for any new words, including the A-words. Mediated generalization to the A-words would then be the number of responses made over and above this base level. This difference is usually called the generalization-index.

The A-zero words, those A-words in a test-list having no L-words in the learning-list to elicit them, were thought to be equivalent to Y-words, and it was expected that the generalization-index to A-zero words would be comparable to that for Y-words. In fact, the mean number of responses to the A-zero words was 0.34, while to the Y-words it was only 0.12, a significant difference, $t\ (160) = 4.28$, $p < 0.01$. Following the experiment, additional data from free association were gathered, which indicated that there existed some associative strength between a few L-words and A-words which had not been identified earlier in the experiment. Though this provided a partial explanation, much of the difference remains unexplained.

The presence of this unexplained effect leads to an important point about the experimental controls. The marked experimental effects uncovered for Lists, Randomizations and Y-words underscore the value of such controls when investigating words. The large number of dimensions on which words vary and the complexity of their interrelationships make it risky to draw conclusions from a small sample or from a single arrangement of items.

SUMMARY

The concept of an implicit response with stimulus-consequences plays a prominent role in accounts of verbal behavior stemming from association-theory. The concept's effectiveness has been demonstrated in several laboratory situations (*e.g.* paired-associate learning, mediated generalization). Current interest centers on the variables that affect the contribution of implicit responses in these tasks. One such variable is the number of elicitations of a specific implicit response in the learning-stage of a motor-response generalization task patterned after that of Mink.

Twelve sets of words were selected from associative norms such that the three stimulus-words in each set elicited a common response. From these stimulus-words learning-lists were so constructed that a specific implicit response would be aroused zero, one, two, or three times. *S*s learned to make a motor-response to each item of the learning-list. Separate lists were so compiled that words would be counterbalanced across conditions. A corresponding test-list was constructed for each learning-list containing all learning-words plus the 12 response-words and appropriate control words. *S* was instructed to respond to words he had pressed to before. Forty *S*s were used with standard mediated generalization-procedures and instructions.

The number of "false presses" to response-words in the test-list indicated the operation of the implicit occurrence of those response words in the learning-stage. The hypothesis was that the number of "false presses" would be a direct function of the number of stimulus-words converging on an implicit response.

The mean number of "false presses" in each category per test-list randomization was: 0—0.34; 1—0.42; 2—0.64; 3—0.92. A trend analysis showed a significant linear trend, $F\ (1, 36) = 30.23$, $p < 0.001$, although the trend was not present in each of the four test-lists. Significant main effects for the control factors further indicated the complexity of the task and the difficulty of working with natural language materials. It is clear, however, that the amount of generalization of a motor-response can be predicted from the number of elicitations of specific implicit responses.

THE PREDICTION OF INDIVIDUAL ASSOCIATION-HIERARCHIES FROM CULTURAL FREQUENCIES

By ALBERT SILVERSTEIN, University of Rhode Island [*]

The use of free-association norms to describe verbal materials has been based on the assumption that "the distribution of responses taken one per individual per stimulus word . . . is characteristic of the distribution of associations at different times in any one representative individual."[1] The assumption seems reasonable, but has not gone completely unchallenged,[2] and the use of averaged values to represent individual verbal processes recently has come in for some very heavy criticism.[3] A direct comparison of cultural frequencies with individual frequencies is badly needed.

There is some evidence that suggests that cultural frequencies do predict individual association-strengths. Rosen and Russell found that the mean cultural frequencies of the first of two successive associations to 100 Kent-Rosanoff stimuli were substantially higher than those of the second.[4] Garskoff found that the median correlation between emission-order of continuous associates and cultural frequency of single associates for 20 words was 0.79.[5] Garskoff and Marshall found a correlation of 0.76 between measures of the associative overlap for 140 words obtained from single and continuous associations.[6] These data were obtained by the production-method, which requires that *S* respond repeatedly to each stimulus-word during a specified period. In this method, *S* cannot report his associations to a stimulus and the only possible measure of associative strength is the *order* of emission, a less direct measure than probability of repetition. A more serious problem lies in the scorer's inability to eliminate associative chains from the data, with the result that successive responses are likely to be associates to prior responses.[7]

---

[*] Received for publication May 9, 1966. This research was supported in part by University of Rhode Island Research Grant-in-Aid No. 132–41.

[1] James Deese, On the structure of associative meaning, *Psychol. Rev.*, 69, 1962, 165.

[2] See, for example, B. F. Skinner, *Verbal Behavior*, 1957, 27, 74; W. A. Russell, Assessment versus experimental acquisition of verbal habits, in C. N. Cofer, (ed.), *Verbal Learning and Verbal Behavior*, 1961, 110–123.

[3] W. K. Estes, Learning theory and the new mental chemistry, *Psychol. Rev.*, 67, 1960, 207–223.

[4] Ephraim Rosen and W. A. Russell, Frequency-characteristics of successive word associations, this JOURNAL, 70, 1957, 120–122.

[5] B. Garskoff, The relation between single word-association and continued association hierarchies, *Psychol. Rep.*, 16, 1965, 307–309.

[6] B. Garskoff and G. R. Marshall, Relations between two measures of verbal relatedness, *Psychol. Rep.*, 16, 1965, 17–18.

[7] Indirect evidence on the relation in question that escapes these criticisms comes from Harold Schlosberg and Charles Heineman, (The relationship between two measures of response strength, *J. exp. Psychol.*, 40, 1950, 235–247), who found a correlation of −0.80 between cultural frequency and associative reaction-time.

A more appropriate method for comparing cultural associative norms with individual hierarchies would be that of independent repetitions, in which *S* is tested with the entire list of stimuli on several occasions. The method of independent repetitions does involve the problem that prior items in the list may influence the response to any given stimulus,[8] but the ability of a response to withstand the effects of shifting context from test to test should be an index of its associative strength. For this reason, it might be best to compare responses of differing cultural frequency with regard to the probability of repetition *after having been given on the first trial.* The only available data on the relation of cultural norms to individual hierarchies obtained by the method of independent repetitions come from a clinical setting. Laffal found that *S*s who failed to repeat an association to a stimulus were more likely to have given a response that had strong cultural competitors than were *S*s who did repeat an association.[9] The present study was designed to obtain more systematic information on this relation by showing *S*s a list of words three times and calculating the frequency with which first-trial associations were repeated. It was hypothesized that this frequency would increase with the cultural frequency of the association. Different groups of *S*s were given word-lists of high and low frequency. Since the former show greater communality of associates, it was hypothesized that they would produce a higher level of repetition.

METHOD

Subjects and apparatus. The *S*s were 96 college students from an introductory course in psychology. Half were assigned at random to each of two list-conditions: high frequency and low frequency. The *S*s were tested in small groups ranging in number from 1–5 with a mean of 3.4. A Bell and Howell Robomatic slide-projector with automatic timer showed each word for 5 sec. on a milk-glass screen.

Materials. Two lists of two-syllable nouns were used, one at each frequency-level. Each list contained 20 words. The Thorndike-Lorge L-count for the High-F nouns ranged from 1000–3300 occurrences per million, while for the Low-F nouns it ranged from 10–33 occurrences.[10] Familiarity, interitem associative strength, and rated meaningfulness co-varied with frequency, but the lists were equated for formal similarity. Association-norms for these words had been obtained previously from 1000 *S*s at the University of California.

Procedure. The *S*s in each group were shown their list three times. The intertrial interval was 25 sec. All *S*s were instructed that they were participating in an experiment on the consistency of association. They were asked to cover the association given to each word before they saw the next stimulus-word, and told that the list of words would be shown in a different order on each presentation. They were asked *not to try* to recall previous associations, but simply to give the first word that came to mind, regardless of what it was. Four different random orders of presentation were used for each of the lists, and 12 *S*s in each group had each of the orders first. Four *Ss* in the Low-F group and two in the High-F group were discarded because they had served recently in another experiment on verbal learning.

---

[8] D. H. Howes and C. E. Osgood, On the combination of associative probabilities in linguistic contexts, this JOURNAL, 67, 1954, 241–258.

[9] Julius Laffal, Response faults in word association as a function of response entropy, *J. abnorm. soc. Psychol.*, 50, 1955, 265–270.

[10] E. L. Thorndike and Irving Lorge, *The Teacher's Wordbook of 30,000 Words*, 1944.

RESULTS

Two tests indicated that the present sample of *Ss* was quite representative of the 1000 *Ss* used for the California association-norms. First, the distributions of percentage of responses falling in each of the first five California ranks were very similar for the two samples. Indeed, the major difference was that the present sample of *Ss* gave a *greater* percentage of California-norm primary responses than did the *Ss* from whom those norms were obtained. Secondly, a comparison was made of the five responses given most frequently by both samples to each of the 40 stimulus-words. In only 80 instances of a possible 400 did any word appear in the first five ranks of one sample, but not the other. For 31 of the 40 stimulus-words, the same primary response was given by both samples.

Each association given on the first trial of this experiment was classified according to whether it fell in one of the first five frequency-ranks of the California norms, in ranks above five, or was idiosyncratic. Then the associations were assigned to one of three repetitive categories: (1) repeated on both subsequent trials; (2) repeated on one of the subsequent trials;[11] or (3) repeated on neither of the subsequent trials. The percentage of responses from each of the seven cultural-frequency categories that were repeated twice, once, or not at all was then calculated. The upper half of Table I presents these data for the pooled responses from both lists, as well as the number of responses obtained at each frequency-level. It is apparent that the probability of obtaining two repetitions is directly related to the relative frequency of the association, while the probability of obtaining no repetitions is inversely related to relative frequency. Both functions show five discrete categories (with Ranks 3, 4, and 5 forming a single category). The probability of obtaining one repetition does not vary in any systematic way. Chi-square tests were performed separately on the distributions obtained from the high- and low-frequency lists to test the hypothesis that the cultural frequency of the responses did not change the distributions across the three repetition-categories. The expected values used for these tests were calculated by taking the percentage of all responses to a list that fell in each of the repetition-categories and multiplying it by the number of responses at each cultural-frequency rank. The values obtained were 101.29 and 71.51 for the High-F and Low-F lists, respectively ($df = 20$, $p < 0.001$, for both).

---

[11] This category was originally divided into: (a) repeated on Trial 2, and (b) repeated on Trial 3. Since the frequencies in these two categories were nearly identical in all cases, they were combined.

The first associations given were also classified into seven percentages of absolute cultural frequency to see whether probability of repetition is more sensitive to this variable than to relative frequency (*i.e.* rank).[12] As may be seen in the lower half of Table I, this distribution of repetitions was substantially the same as that for using rank. Again, five discrete categories were found: (1) over 25.0%; (2) 5.0%–24.9%; (3) 0.6%–4.9%; (4) 0.2%–0.5%, and (5) idiosyncratic.

TABLE I

PERCENTAGE OF ASSOCIATIONS REPEATED IN BOTH LISTS AS A FUNCTION OF CULTURAL FREQUENCY

A. By Frequency-rank

Rank	No.	%2 Repeats	%1 Repeat	%0 Repeats
1	543	53.6	24.7	21.7
2	213	44.1	28.7	27.2
3	133	39.8	24.1	36.1
4	107	36.5	32.7	30.8
5	60	40.0	23.3	36.7
>5	506	29.0	28.1	42.9
Idiosyncratic	217	19.4	19.4	61.3

B. By Frequency-percentage

Frequency	No.	%2 Repeats	%1 Repeat	%0 Repeats
>25.0	356	58.4	24.2	17.4
12.0–24.9	323	44.3	27.9	27.9
5.0–11.9	271	42.4	27.7	29.9
2.0– 4.9	249	31.7	26.9	41.8
0.6– 1.9	218	33.0	26.1	40.8
0.2– 0.5	145	24.8	26.3	48.9
Idiosyncratic	217	19.4	19.4	61.3

To investigate further the influence on repetitive probability of absolute and relative frequency, an index of repetition was calculated for the five primary associates in each list with the lowest absolute frequencies, five Rank-3 associates in each list with matching absolute frequencies, and the five primary associates with the highest absolute frequencies. The index of repetition (*IR*) represents the total number of times that an associate was repeated on the second and third trials divided by the total number of opportunities. The most common primary responses were repeated more often than the least common primaries (71.6% *vs.* 59.3%, $t = 1.75$, $df = 19$, $0.05 < p < 0.10$), while the *IR*s for primaries and Rank-3 responses of matched absolute frequency did not substantially differ (59.3% *vs.* 59.8%, $t < 1$).

---

[12] These frequency-categories were empirically derived from 'natural' cut-off points in the frequencies with which different frequency-values were obtained.

The hypothesis that high-frequency stimuli would produce greater associative repetition than low-frequency stimuli was not confirmed, despite the fact that the High-F list did yield far greater communality of responses. The over-all indices of repetition for the High-F and Low-F lists, respectively, were 52.1% and 51.7%. Apparently, the greater number of infrequent associates given in the Low-F list was offset by their more frequent repetition. For the 93 instances in which once- or twice-given associates appeared in the High-F condition, the *IR* was 23.1%, while for the 185 instances of once- or twice-given associates in the Low-F condition, the IR was 33.8% ($X^2 = 6.49$, $df = 3$, $0.05 < p < 0.10$).

Discussion. The major finding of the present experiment is that the norms of cultural free-association do reflect individual associative hierarchies as measured by probability of repetition. It appears, however, that the associative strengths obtained from the single associations of a large number of *S*s yield a finer gradation than those obtained from probabilities of repetition within *S*s. This fact is in accord with some recent data showing that weaker cultural associations may produce as rapid paired-associates learning, direct and mediated, as do stronger cultural associations.[13] Perhaps cultural-free association-norms produce a spurious fineness of gradation in associative strength. An alternative interpretation suggested by Storms is that 'recency' may have operated to strengthen temporarily the weaker associations.[14] He showed that, if the *S*s are given a list of words to recall prior to an associative task, the words of the list are more likely to appear in the protocols of association. It would be useful to see if the gradation in repetitive probability could be sharpened by using very long intervals between trials.

The present data strongly suggest that absolute cultural frequencies are more reliable indices of intra-*S* strengths of association than are cultural-frequency ranks. Primary and Rank-3 responses of equal absolute frequency showed equivalent *IR*-values, while the same primary responses were repeated 12.3% less often than more commonly given primaries. It is possible that some factor other than absolute associative frequency is responsible for this difference. For example, primary associates of greater frequency have greater superiority over their nearest competitor than do primary associates of lower frequency. Furthermore, common primaries might be more likely to occur to stimuli that produce a smaller total number of different associations, or to stimuli that produce more associations that are related to the primary associate. Although it is not possible to decide among these alternatives with the small number of stimuli used in this experiment, it would seem preferable to use the frequencies rather than the ranks from cultural norms of association for diagnostic purposes.

---

[13] J. J. Jenkins, Mediated associations: Paradigms and situation, in C. N. Cofer and B. S. Musgrave (eds.) *Verbal Behavior and Learning: Problems and Processes,* 1963, 210–245; The 1952 word association norms, in Leo Postman (ed.), *Norms of Word Associations* (in press).

[14] L. H. Storms, Backward association in verbally mediated learning. Unpublished Doctoral dissertation, University of Minnesota, 1956.

The lack of a difference in the over-all level of repetition for the two lists was shown to be due mainly to the higher level of repetition for infrequent responses in the Low-F list. It may well be that idiosyncratic responses represent different things for frequent and infrequent stimuli. For one thing, High-F words are likely to have been experienced in a wider variety of linguistic contexts by any one individual than have Low-F words, while the dominant linguistic context for Low-F words should be more variable among individuals. Thus, an idiosyncratic response to a High-F stimulus may represent the influence of some transient feature of the momentary context, while an idiosyncratic response to a Low-F stimulus may be a truly strong associate for that individual which he does not share with other individuals. For these same reasons, one would expect idiosyncratic responses to High-F stimuli to be subject to a greater number of competing associations than are idiosyncratic responses to Low-F stimuli. Such an interpretation is supported by the fact that High-F words yield a greater number of different associations per person by the production-method,[15] despite the fact that they show less associative variety in single-response hierarchies. It is thus quite possible that cultural association-norms give a somewhat inaccurate picture of the individual response-hierarchies for extremely unfamiliar words.

It is also possible that the level of repetition for more common associates of the High-F list was depressed by a general tendency to inhibit repetition of a response to successive stimuli. Thus, if *S* has just said "man" to the stimulus "doctor," he may refrain from saying it again if "woman" is the next stimulus. Since overlap in associative hierarchies is much greater for High-F words than for Low-F words, a reliable difference between *IR*-values for the two lists might have been obtained if they had been mixed in a single list.

---

[15] C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421–430.

SUMMARY

Two groups of *S*s were given three tests of successive free-association with a list of 20 stimulus-words. The words were of high frequency for one group and of low frequency for the second. The index of repetition of associations (*IR*) was analyzed as a function of their popularity in terms of cultural associative norms. The results showed that cultural frequencies do, in general, predict *IR* values, from which it may be inferred that they represent individual associative hierarchies. The cultural associative frequencies proved, however, to be more finely graded than the *IR* values. The only difference between the *IR*s for High-F and Low-F lists was a higher value for idiosyncratic and twice-given responses in the Low-F case.

PERFORMANCE IN A RUNWAY

By JOHN R. STABLER, Louisiana State University, and ANTHONY TAYLOR,
New York State University, Cortland

In his book *Behavior Theory and Conditioning,* Spence relates the principles of classical conditioning to instrumental learning by use of the concept of fractional anticipatory goal-response (r_g).[1] The essential notion involved is that instrumental learning involves both classical conditioning of portions of the consummatory response to stimulus-cues throughout the behavior-chain as well as conditioning of the instrumental acts themselves. In a runway to a goal-box, stimuli from the alley are said to acquire habit-loadings for the 'instrumental' locomoting response both from direct associations—strengthened on the basis of contiguity of the alley-stimuli and the locomotor response; and from indirect associations—strengthened, by stimulus-generalization, by events which occur in the goal-box. Stimuli in the alley which are similar to the stimuli present in the goal-box during the consummatory response become, through the mediation of r_g, conditioned stimuli for the instrumental response.

The Hull-Spence theory assumes that drive and habits, including generalized habits, combine mltiplicatively. Despite the possibility of such a logical inconsistency, a review of the experimental literature on the combination of drive (D) and incentive (K) indicates that the additive hypothesis has received the bulk of empirical support.[2] The present study relates to the theoretical questions and was designed to determine the nature of the combination between D and K in a straight alley. The manipulation of K rested on the Spencerian assumptions that K is directly related to r_g–s_g, and that, in turn, the strength of r_g in the alley is a positive function of the stimulus-similarity between the alley and the goal-box.[3]

Two experiments were conducted. The second was largely a replication of the first.

METHOD AND PROCEDURE

Subjects. The *Ss* in each study were 60 naïve male rats of the Wistar strain obtained from Harlan Industries, Cumberland, Indiana. The rats were 90 days old at the onset of adaptation-procedures. In both studies, two rats were discarded be-

---

* Received for publication September 24, 1965. This research was supported by Research Grant, 15.1.1 from the Division of Research, State Department of Hospitals, State of Louisiana.

[1] K. W. Spence, *Behavior Theory and Conditioning,* 1956, 37–51.

[2] R. W. Black, On the combination of drive and incentive motivation, *Psychol. Rev.,* 72, 1965, 310–318.

[3] Spence, *op. cit.,* 135.

cause of illness or pregnancy, and their scores were prorated from the scores of the remaining animals in their respective groups.

Apparatus. The apparatus for both experiments consisted of two, dimensionally-similar, straight runways. The start- and goal-boxes were 9 in. long, 11 in. wide and 8¾ in. deep. In Experiment I, the walls and floors of the start-boxes were painted flat black and were covered by clear glass. Both goal-boxes, in contrast, were painted flat white, the glass covers were striped at 2 in. intervals with ½-in. wide black tape, and the floors were studded with white thumbtacks placed 3 to 5 in. apart. Each goal-box contained a white food cup located 1 in. from the center of the rear wall. Traversing one alley involved changes from black start-box to black runway to white goal-box (B-B-W); and traversal of the other alley involved changes from black start-box to white runway to white goal-box (B-W-W). The two runways were set up parallel to one another, and 3 ft. apart. Hanging midway between, and about 6 ft. above them. was a string of seven, evenly-spaced 150-w. bulbs. A large strip of aluminum foil was suspended over each alley in such a way as to reflect the light down into the alley.

Two guillotine doors in tandem separated each start-box and its runway. The door on the side of the start-box was opaque fiber board; on the other side, translucent plastic. Suspended at the entrance of each goal-box was a pair of slightly overlapping opaque curtains. The brightness of the sides of both the opaque door and the curtains were consistent with the alley on which they fronted. The starting time was measured with a Standard Electric timer that was activated by the raising of the translucent start-box door, and was stopped by interruption of a photoelectric beam located 0.42 ft. into the runway. Runway-time was recorded with a second timer which started with the breaking of the first photoelectric beam and stopped with the interruption of a second photoelectric beam located 5.33 ft. beyond the first beam. In Experiment II, a third photoelectric beam, going directly over the food cup, and a third timer were added to obtain measures of the final approach to the food cup.

Changes in brightness in the alleys during the first experiment may have served to disrupt the instrumental response; therefore two modifications were introduced in Experiment II. The start-box in the alley with the white runway was painted white. This modification eliminated the brightness change for *Ss* leaving the start-box and increased the similarity of the start-box to the goal-box. One alley thus had a black start-box, a black runway and a white goal-box (B-B-W), and the other alley had white in all three segments (W-W-W-). The tapes on the covers to the alleys and the tacks inbedded in the floors were also varied in a way appropriate to the alley and the goal-box. The second modification reduced the contrast between black and white alleys. Narrow strips of black tape were applied vertically on all white surfaces, and narrow strips of white tape were applied horizontally on all black surfaces.

Procedure. Both experiments employed a factorial design with two levels of deprivation of food: 4 hr. (D_L) and 20 hr. (D_H); and two levels of similarity between alleys and goal-boxes, low similarity (K_L) and high similarity (K_H). In both experiments, 15 animals were assigned to each of the four combined conditions: D_LK_L, D_LK_H, D_HK_L, and D_HK_H.

In both experiments, the animals initially learned to lap sugar water from a cup which was placed in their home cages. They then began 12 days of adaptation to handling and to the food-deprivation schedule associated with their particular experimental assignment. During the six days preceding the first day of training, the *Ss* were adapted to the runway and received four pre-training trials. During the following days, and at the rate of four trials per day, the animals in Experiment I ran 44 acquisition-trials; and the animals in Experiment II ran 64.

The rats were run in two groups. Water was removed from the home cages of both groups 1 hr. before running the experiment began and, depending on the treatment condition, food was removed either 4 or 20 hr. before the first trial. The individual order in which the animals ran the trials was varied daily. After placing the animal in the start-box, and as soon as it oriented toward the alley, the opaque door was raised. The translucent door was raised automatically 0.5 sec. later. The rat was removed from the goal-box immediately after consuming two drops of a 32% sucrose solution which had been previously placed in the food cup. The rat was held in a waiting cage after each trial until the next rat had completed a trial. It was then transferred from the waiting cage to the home cage. The intertrial interval was about 20 min. for each animal. One hour after the last member of a group completed running, all the members received water and were fed *ad lib* in the home cages.

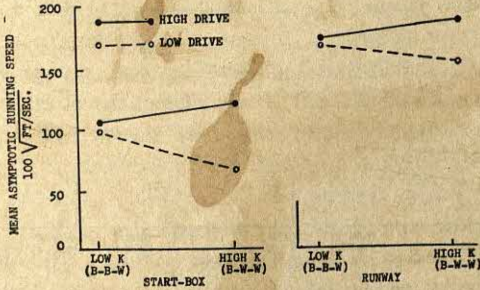


FIG. 1. EXPERIMENT I: RUNNING SPEED OF RATS IN START-BOX AND RUNWAY

Effect of deprivation (D) and similarity of alley and goal-box (K).

RESULTS

Experiment I. The individual mean speeds for start-box and runway segments were calculated for the final 16 trials (Trials 29-44). The individual mean scores for both measures were transformed by taking their square roots in order to increase homogeneity of variance and reduce thereby the correlation between means and variances. The group averages for the transformed start-box and runway measures are plotted in Fig. 1. As can be seen, there is a divergence between the two levels of drive (D) that become larger at a higher level of incentive, that is, with the greater similarity of alley and goal-box. This is suggestive of an interaction effect.

Analysis of variance of start-box data showed a reliable effect of D

($F = 9.80$, $p < 0.01$) and D × K ($F = 6.60$, $p < 0.05$); analysis of variance of the speed in the runway showed, however, only a reliable influence of D ($F = 6.24$, $p < 0.05$). Duncan's range-test showed that Group D_LK_H ran reliably slower than any of the other groups in the start-box segment and reliably slower than Group D_HK_H in the runway-segment.[4]

Experiment II. In the second experiment, individual mean scores for start-box, runway, and goal-box speeds were obtained for the final 16 trials (Trials 49-64). The scores were transformed in the same fashion as the scores in the initial experiment. A graph of the group averages of the transformed scores, as seen in Fig. 2, shows a contrast with the findings of Experiment I. The appearance of an interaction between D and K is no longer present. Also, instead of a downward trend for the low drive condition as a function of K, there is now an upward trend, if any, with speed of running increasing as the alley is more similar to the goal-box.

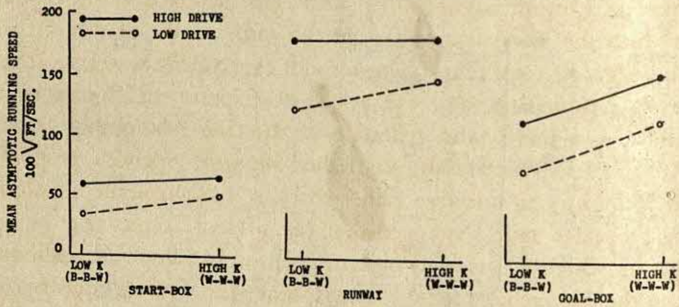


FIG. 2. EXPERIMENT II: RUNNING SPEED OF RATS IN START-BOX, RUNWAY, AND GOAL-BOX

Effect of deprivation (D) and similarity of alley and goal-box (K).

The statistical results also indicate a lack of interaction between D and K. The effect of drive was statistically significant ($p < 0.01$) on all three response measures ($F = 17.84$; $F = 28.19$; $F = 28.58$, respectively), and the effect of K was statistically significant for start-box and goal-box measures ($F = 4.20$, p 0.05; $F = 19.68$, $p < 0.01$). The Duncan range-test showed that Group D_LK_L ran reliably slower than any of the other groups in the start-box and goal-box, and reliably slower than groups D_HK_L and D_HK_H in the runway. Group D_HK_H ran faster than Group D_LK_L in the start-box, faster than Groups D_LK_L and D_LK_H in the runway, and faster than all other groups in the goal-box.

---

[4] D. B. Duncan, Multiple range and multiple F-test, *Biometrics*, 11, 1955, 1–42.

DISCUSSION

Although the data from the start-box of Experiment I indicated the presence of an interaction between drive and incentive, the finding that Group D_LK_H ran slower than Group D_LK_L suggested the possibility that the interaction-effect was an artifact of experimental procedures. The finding of a statistically significant interaction in the start-box but a less significant one in the runway further prompted such a conclusion. It was thought that the inhibiting consequences of moving from a black start-box into a white runway in the high K condition interfered with the facilatory effect of the fractional anticipatory goal-response. Animals running from the start-box into the runway experienced an abrupt stimulus-change, from a black box in a white alley, and it is known that rats prefer a dark environment to a bright environment. Such considerations led to Experiment II, in which the alley-segments were the same as the start-box for both low and high K conditions. The second experiment was free from this source of confounding, and the results of Experiment II may be taken, we believe, as a more valid reflection of the true relationship between D and K variables. The present study thus supports Spence's proposal that D and K join in an additive rather than in a multiplicative fashion.

It is proper to note, however, that the present results may be specific to instances where K is varied by manipulation of alley and goal-box similarity. Although such a method is consistent with Spence's theory of fractional anticipatory goal-responses, it is a less direct procedure than the procedure of varying the magnitude of the incentive object in the goal-box. Furthermore, the feeding schedule induced both a drive effect and differential weight gain between low and high drive animals. Another feeding procedure, one which controlled for weight gain, might possibly lead to different results. A review of the literature on drive and incentive combinations illustrates, as does the present study, that the particular nature of the motivating and reinforcing operations affects the behavioral outcome.[5]

SUMMARY

The present study was concerned with the nature of the variables of drive (D) and incentive (K) when incentive is varied by manipulation of the similarity of alley and goal-box in a straight runway situation. Although the starting time in leaving the start-box in the first experiment suggested that D and K interact, a subsequent replication, which corrected for possible confounding from changes in brightness from start-box to the alley, showed no D × K interaction. The results were interpreted as support for Spence's additive hypothesis.

---

[5] Black, *op. cit.,* 310–318.

EFFECTS OF SERIAL PRESENTATION OF PAIRED ASSOCIATES ON PERFORMANCE DURING AND AFTER TRAINING

By SLATER E. NEWMAN, North Carolina State University

McGeoch and McKinney compared the effects of serial and random presentation of paired associates.[1] They found that when pairs were presented in the same serial order on five consecutive training trials the *Ss* performed better on an immediately-subsequent test than when the pairs were presented in a different order on each training trial. The present experiment was similar to theirs except that (1) a test followed each of the training trials; (2) both the Serial and Random groups had the same random-order test after training; and (3) tactual stimuli were used so that *S*'s behavior could be observed as he explored these stimuli. Before training began, the *Ss* in both groups were told about the order in which the items would appear during training (as in the McGeoch and McKinney experiment) and on the final test.

METHOD

Subjects. The *Ss* ($N = 40$) were undergraduates at North Carolina State University. All were right-handed. They were randomly assigned with the restriction that each group would have five men and five women.

Materials. The stimulus-terms were six-sided closed random shapes made of plastic, approximately ½ mm. thick. Method 1 described by Attneave and Arnoult was used in determining the shapes.[2] The widths ranged from 1.50 to 3.50 in. and the heights from 1.25 to 2.50 in. The response-terms were the consonants J, L, Q, R, X, Z.

The shapes were mounted on a 'tactual memory-drum.' This 'drum' was made by mounting a phonograph turntable on a desk. A plywood disk, 15.5 in. in diameter was centered on the turntable and was covered with window-shade cloth. The base of each shape was approximately 1 in. from the edge of the turntable. The shapes were separated from each other on the turntable by about 4 in. A seventh shape was located in the center of the turntable and was used for practice.

---

* Received for publication December 10, 1965. From the Institute of Human Learning, University of California, Berkeley. This research was supported by Contract Nonr 486(08) with the Office of Naval Research. The assistance of Mrs. Leona Markman, who collected the data, and of Mr. Charles Williams, who designed and constructed the apparatus, is gratefully acknowledged.

[1] J. A. McGeoch and Fred McKinney, Studies of retroactive inhibition: VIII. The influence of the relative order of presentation of original and interpolated paired associates, *J. exp. Psychol.*, 20, 1937, 60–83.

[2] Fred Attneave and M. D. Arnoult, The quantitative study of shape and pattern perception, *Psychol. Bull.*, 53, 1956, 452–471.

Procedure. After the *S* was seated at the desk, *E* read the instructions to him. *S* was told about the nature of the task and about the order in which the items would appear during training and on the final test which was immediately to follow training. *S* was then blindfolded and his right index finger was placed on a wooden platform in front of him. The top of the platform was 3/16 in. from the edge of the turntable.

Shortly after this, *S*'s finger was placed at the lower left corner of the practice shape and he was given 10 sec. to trace its outside circumference. When *E* said "Stop," *S* returned his finger to the platform to await the beginning of training.

A procedure alternating a paired training trial with a test was used. As each shape was presented on a training trial, *S* was told its name and was given 10 sec. to explore its outside circumference. (A brake operated by *E* kept the turntable from moving.) At the end of 10 sec., *E* said "Stop." *S* then returned his finger to the platform and waited 7 sec. for the next shape. On all tests *S* was also given 10 sec. to explore each shape and to call out its name. Again there was a 7-sec. interval between items.

Half the *S*s were presented with the items in the same order on each training and test trial; for the other *S*s the items were presented in a different order on each trial. For half the *S*s in the Serial group the items were presented in one order, A, for the rest in a second order, B. Two sets, each of four different orders, were used for the Random group; half the group was exposed to each set.

After the second test all *S*s were given a third test in which the order was different from any previously used with any group. The rate of presentation on this test was the same as on Tests 1 and 2. Finally each *S* was tested for recall of the response terms above; he was given 10 sec. to call out all of the terms he could remember.

A tape-recorder was used to program the experiment through the end of Test 3. Thus *E* was free to keep a written record of *S*'s finger movements while exploring the shapes and of his verbal responses on all tests.

RESULTS

Tests 1 and 2. Since there were no differences between the sub-groups receiving the same treatment, the results for both sub-groups were combined. Table I presents the means for six measures on Tests 1, 2, 3, and for three measures on recall.

The number of correct responses on each trial is the most direct measure of performance. With this measure the Serial group does better than the Random group in every instance. Moreover, the superiority of the Serial group is greater both in absolute and in relative terms on the second trial. As might be expected, both groups are better on Test 2.

An analysis of variance appropriate for repeated measurements was done for correct responses. The *F*s for Treatments, Trials, and interaction, given in Table II, were all significant ($p < 0.01$). The significant interaction was due to the greater difference between the Serial and Random means on Test 2. The F-tests that result from this analysis are in the second column of Table II.

Results from the other measures show that the Serial group made more different responses than the Random group and both groups made more different responses on the second test. The Serial group made fewer repetitions using responses from the list.

The reliability of these differences was tested by similar analyses, and the Fs are given in Table II. The F for interaction was not significant ($p > 0.05$) for any of these variables while the F for Trials was significant ($p < 0.05$) for all but repeated substitutions. The F for Treatments was significant ($p < 0.05$) for different responses and for repetitions.

TABLE I

PERFORMANCE ON EACH TEST IN MEAN NUMBER OF RESPONSES

Dependent variable	Group	Test			Recall
		1	2	3	
Correct	Serial	2.20	4.40	3.40	5.15
	Random	1.65	2.65	2.35	5.00
Different	Serial	4.05	5.30	4.65	
	Random	3.50	4.45	4.55	
Inventions	Serial	.35	.20	.15	.30
	Random	.45	.05	.30	.30
Omissions	Serial	1.35	.30	.20	
	Random	1.45	.95	.45	
Substitutions (repetitions)	Serial	.25	.20	1.00	.15*
	Random	.60	.55	.70	.05*
Substitutions (nonrepetitions)	Serial	1.85	.90	1.25	
	Random	1.85	1.80	2.20	

* For recall this is the number of responses repeated.

Tracing behavior during training was also studied. The results appear in Table III. The measure was the number of times each point of a shape was touched. The Ss touched more points on training trials than on test trials, and more on the second training and test series than on the first. The Serial group did not differ, however, from the Random group in this respect nor was the interaction between training and trials significant ($p > 0.05$). The F for Type of Trial was 7.41 (1/114, $p < 0.01$) and for Stage of Training 32.45 (1/114, $p < 0.01$).

Test 3. As may be seen in Table I, the superiority of the Serial group remained in Test 3. T-tests were done for the same six variables analyzed for Tests 1 and 2. For correct responses, t was 2.23 ($df = 38$, $p < 0.05$). The Serial group made more. The only other significant effect was for nonrepeated substitutions. The Random mean was higher ($t = 2.16$, $df = 38$, $p < 0.05$).

A repeated measurements analysis of variance was done for tracing behavior on Tests 2 and 3. Only the effect of Trials was significant ($F = 4.45$, $1/38$, $p < 0.05$). Fewer points were touched on Test 3 than on Test 2.

Recall of the response-letters. T-tests showed that the means for the Serial and Random groups did not differ on any of the variables ($p > 0.05$) on recall. The means appear in Table I.

To determine whether Serial *S*s may have learned the order of the responses, rank-order correlations were done for every *S* between the order in which the responses appeared during training and the order in which *S* gave them during recall. Similar correlations were done for the Random *S*s involving the order of responses during recall and order A or B (whichever gave the higher correlation for each Random *S*). A Kruskal-Wallis test was then applied to these data. The correlations were higher for the Serial group ($H = 4.62$, $df = 1$, $p < 0.05$).

TABLE II

F-TESTS ON SIX VARIABLES: TEST 1 AND 2. (DEGREES OF FREEDOM ARE 1 AND 38)

	Source		
Variable	Treatments	Trials	Interaction
Correct responses	8.62†	83.93†	43.44†
Different responses	8.52†	33.61†	.72
Omissions	2.20	12.02†	1.90
Inventions	.06	6.61*	1.30
Repeated substitutions	7.21*	.16	.00
Non-repeated substitutions	2.16	4.13*	.00

* $p<.05$.
† $p<.01$.

TABLE III

MEAN NUMBER OF POINTS TOUCHED ON EACH TRIAL

Group	Training Trials		Test Trials		
	1st	*2nd*	*1st*	*2nd*	*3rd*
Serial	60.10	79.00	56.80	67.15	66.65
Random	62.25	75.80	59.65	68.75	63.95

DISCUSSION

The results for performance during training show that (compared with Random Presentation) Serial Presentation led to more correct responses (just as in the McGeoch and McKinney experiment) and to more different within-list responses and fewer repeated substitutions. The results for free recall of the responses suggest that this facilitation during training may have resulted, at least in part, from *Ss'* learning the order of the response terms. Data from Newman are similarly interpretable.[3]

Of particular interest is that Serial Presentation facilitated performance on the random-order test which followed training. Possible explanations for this are that serial presentation facilitated discrimination of the tactual stimuli or associative learning. Although learning of the responses could also have been facilitated, such an explanation does not seem tenable, since the Random Group gave approximately the same number of within-list responses as the Serial Group on Test 3, and also during free recall of the responses.

In several recent experiments, *Ss* were not told beforehand about the order in which the items would appear during or after training. The results from these experiments have not been consistent. Though facilitation was observed by Newman and Saltz and by Newman, it was not observed in the two experiments of Martin and Saltz.[4] The results of these experiments combined with those reported here and with McGeoch and McKinney's suggest that serial presentation is likely to facilitate performance during training if *S* is told prior to training that the items will be presented serially. If *S* is not given this information, facilitation is less likely. The results of these experiments suggest further that performance following training will be facilitated if, in addition, *S* is told prior to training about the order of the items on the post-training test. Future experiments will be necessary to test the validity of these hypotheses.

SUMMARY

Two groups of blindfolded *Ss* were trained on a list of six paired associates. The stimulus-terms were random shapes presented on a 'tactual memory-drum'; the response terms were consonants presented auditorily. A Serial group was exposed to the items in the same order on the two training trials and on the two test-trials; for the Random group a different order was used on each of the training and test-trials. The Serial group gave more correct responses than the Random group both during training ($p < 0.01$) and on a random-order test which followed ($p < 0.05$).

---

[3] S. E. Newman, Serial position as a cue in learning: The effect of test rate, *J. exp. Psychol.*, 76, 1966, 319–320.

[4] Newman, *op. cit.*, 319; S. E. Newman and Eli Saltz, Serial position as a cue in learning, this JOURNAL, 75, 1962, 102–108; C. J. Martin and Eli Saltz, Serial vs. random presentation of paired associates, *J. exp. Psychol.*, 65, 1963, 609–615.

THE RELATION BETWEEN AGE AND THE MAGNITUDE OF THE PONZO ILLUSION

By H. W. LEIBOWITZ and J. M. JUDISCH, The Pennsylvania State University

Geometric-optical illusions have occupied a prominent place in the visual and perceptual literature, but they usually have been treated as interesting demonstrations or 'oddities,' with few attempts to analyze quantitatively the parameters which influence their magnitude or the mechanisms which are responsible for them.[1] It has been demonstrated that certain visual illusions vary systematically as a function of age.[2] In particular, the Ponzo illusion (Fig. 1) has been shown to increase with age from childhood through adolescence, after which it remains stable through the college-years.[3] The purpose of the present study is to confirm and extend these results over as wide an age-range as possible under comparable testing conditions. The *Ss*, ranging in age from childhood (3.5 yr.) to old age (88 yr.), were tested in order to obtain a more complete developmental function and to determine the effect of aging on the illusion.

Apparatus. The converging lines of the illusion were constructed by arranging black "graph-a-plan" tape, ⅛ in. wide, as indicated by the pattern in Fig. 1. Twenty-one 11″ × 14″ reductions of this pattern were made photographically (Eastman Kodak Polycontrast G exposed through a #3½ Bessler filter) and mounted on ¼ in. posterboard. Each reproduction consisted of the array of 11 lines, each 1/16 in. wide, radiating at 10° intervals from a common apex. The length of the center-line is 12 9/16 in. The standard and variable stimulus-lines were produced by applying the same "graph-a-plan" tape to the photographs, which were oriented perpendicular to, and bisected by, the central ray of the bundle. The length of the standard stimulus, located nearer the open end of the array, was constant at 4 in. The variable stimuli, located at the apex end of the array,

---

* Received for publication April 11, 1966. This research was supported by Grant MH-08061 from the Public Health Service.

[1] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 238–245.

[2] Alfred Binet, La mesure des illusions visuelles chez les enfants, *Rev. Phil.*, 40, 1895, 11–25 (transl. by R. H. Pollack and F. K. Zetland, *Percep. Motor Skills*, 20, 1965, 917–930; Jean Piaget, *Les Mécanismes Perceptifs*, 1961, 1–457; Seymour Wapner, Heinz Werner, and P. E. Comalii, Jr., Perception of part-whole relationships in middle and old age, *J. Geront.*, 15, 1960, 412–416; Charles Hanley and D. J. Zerbolio, Developmental changes in five illusions measured by the up-and-down method, *Child Develop.* 36, 1965, 437–452.

[3] H. W. Leibowitz and M. Heisel, L'évolution de l'illusion de Ponzo en function de l'âge, *Arch. Psychol.* (Geneva), 36, 1948, 328–331; Hanley and Zerbolio, *op. cit.*, 437–452.

ranged in length from 2½ to 5 in. in ⅛ in. steps. The variable line was 2½ in., and the comparison-line 10⅝ in., from the apex.

Procedure. The stimuli were displayed horizontally on a table before *S*, who was asked to point or tell which of the two lines, comparison or variable, was "longer" or "bigger." Both terms were used to avoid ambiguity. The length of the variable line which appeared equal in length to the comparison-line was determined individually for each *S* by interpolation depending upon the length of the variable stimulus at which a transition in response occurred. The same randomized order of presentation was used for every *S*. For cases in which the transition was not smooth, the stimuli in the region of the transition were presented again to *S*.

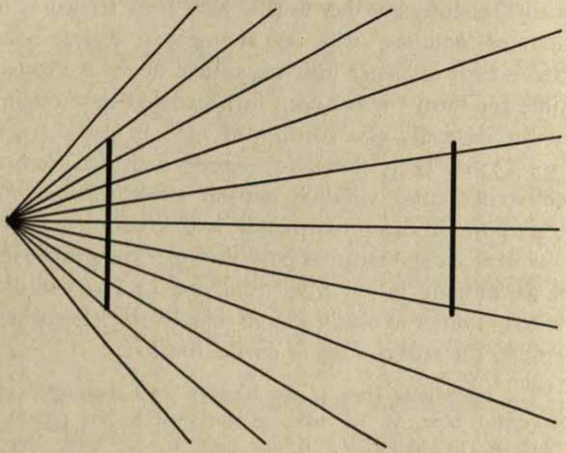


FIG. 1. THE PONZO ILLUSION
The two vertical lines are of equal length.

Since data were obtained in a number of different institutions depending upon the population being tested, it was not possible to maintain precise uniformity of illuminance, which was approximately 50 ft.-C in all testing situations. The viewing distance was approximately 18 in.

Subjects. The 427 *Ss*, obtained from six different institutions, ranged in age from 3.5–88 yr. The preschool and college-age *Ss* were obtained at the University, the school-age children from the local public schools, the 20-47-year-olds from a state prison, and the older *Ss* from the county old-age home.[4] Except for the prisoners, all of whom were men, each group was made up of equal numbers of *Ss* of both sexes. The age-range for each group through the college years was less than one year, while in the older groups the range was somewhat larger.

---

[4] The authors are indebted to the following persons for their coöperation in making *Ss* available: Dr. Winona Morgan, Department of Child Development, The Pennsylvania State University; Mr. Robert Yoho, Miss Pauline Budinger, and Mr. Gerald Wiser, Bald Eagle Area School District; Mr. A. C. Cavell, State Correctional Institution at Rockview; and Mrs. Gerald Kline, Centre Crest Home.

All *Ss* were in 'normal' health. The older *Ss*, as indicated by hospital records and staff-evaluations, were alert, ambulatory or semi-ambulatory, were able to read, and had no record of visual or sensory pathology. It was not possible to select the *Ss* with respect to intelligence, but previous observations indicate that the relation between intelligence-level and the magnitude of the illusion would be negligible in these *Ss*.[5]

TABLE I

THE MAGNITUDE OF THE PONZO ILLUSION AS A FUNCTION OF AGE

The measure of the illusion is the length of the variable line located near the apex of the figure which matched a 4-in. comparison-line located near the open end of the figure. Each group consisted of 10 *S*s except the group of 21.5-year-old men, which contained 7 *S*s.

Males				Females			
Mean age (yr.)	Range	Length of comparison-line (in.)	S.D.	Mean age (Yr.)	Range	Length of comparison-line (in.)	S.D.
4.1	3.5– 4.7	4.09	0.19	4.2	3.5– 4.7	3.86	0.20
5.2	4.9– 5.8	3.90	0.22	5.2	4.9– 5.5	3.91	0.19
6.4	5.9– 6.8	3.81	0.27	6.2	5.9– 6.8	3.63	0.19
7.3	6.8– 7.9	3.69	0.16	7.3	6.9– 8.7	3.57	0.22
8.3	7.9– 8.7	3.58	0.18	8.3	7.9– 8.6	3.60	0.24
9.2	8.8– 9.6	3.62	0.17	9.3	8.9– 9.7	3.60	0.15
10.3	9.9–10.7	3.51	0.29	10.4	9.9–10.8	3.52	0.27
11.1	10.8–11.5	3.66	0.20	11.4	10.9–11.7	3.62	0.22
12.7	12.1–13.8	3.99	0.19	12.4	12.1–17.9	3.37	0.19
13.7	13.2–14.7	3.42	0.27	13.6	13.0–14.1	3.30	0.30
14.7	14.2–15.9	3.64	0.23	14.5	14.3–15.5	3.52	0.19
15.6	15.2–16.4	3.47	0.29	15.4	15.0–15.7	3.50	0.26
16.6	16.2–17.6	3.47	0.12	16.6	16.3–17.7	3.55	0.20
17.7	17.2–18.2	3.56	0.14	17.5	17.0–18.1	3.56	0.26
19.3	17.8–20.3	3.39	0.24	18.1	18.0–19.4	3.51	0.24
21.5	20.6–27.4	3.47	0.20				
23.0	22.5–23.8	3.40	0.28				
24.4	23.6–24.7	3.45	0.15				
25.9	25.1–27.3	3.50	0.26				
28.6	27.5–30.4	3.49	0.29				
31.2	30.5–32.1	3.47	0.27				
36.1	34.0–37.6	3.55	0.16				
40.2	38.6–42.4	3.55	0.16				
45.7	42.7–47.5	3.42	0.19				
61.4	48.9–69.8	3.84	0.20	58.3	44.3–70.0	3.82	0.33
82.5	77.6–88.0	3.84	0.25	82.7	69.7–88.1	3.75	0.33

Results. The results are presented in Table I. Since there were no consistent sex-differences, the results for *Ss* of both sexes have been averaged and plotted as a function of age in Fig. 2. It will be observed that the illusion is essentially nonexistent for the youngest *Ss*. With increasing age, it increases rapidly up to about 13 yr., after which it levels off through age

---

[5] Leibowitz and Heisel, *op. cit.*, 328–331.

50. For the oldest *Ss* tested, there is a marked decrease in the magnitude of the illusion.

Discussion. The increase in the magnitude of the illusion as a function of age from 4–13 yr. confirms previous findings.[6] Furthermore, the stability of the illusion beyond adolescence is seen here to extend to approximately age 50. Of particular interest in the present data is the unexpected decrease in the magnitude of the illusion in the still older *Ss*. In these *Ss*, the illusion is no more marked than it is in 5-yr. old children.

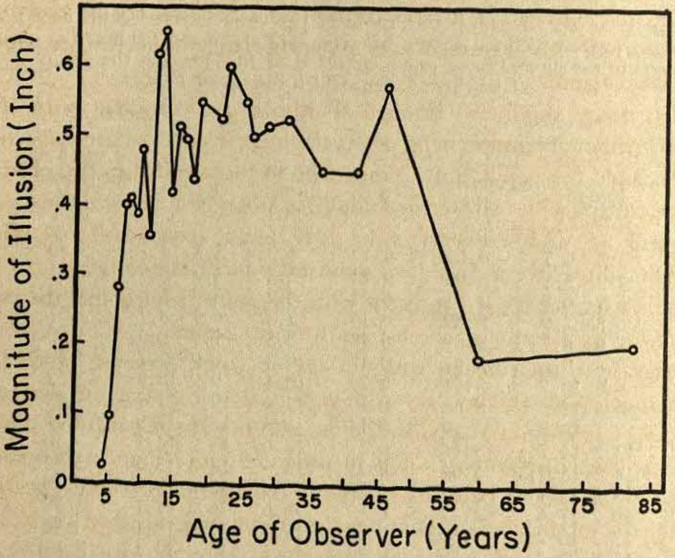


FIG. 2. THE MAGNITUDE OF THE PONZO ILLUSION AS A FUNCTION OF AGE
The magnitude is calculated as the underestimation of the length of the 4-in. standard-line, located near the open end of the figure, as computed from the setting of the variable line.

Theoretical analyses of the Ponzo illusion have been based on the suggestions by Tausch and by Von Holst that this illusion is an example of the operation of a normal mechanism in an inappropriate situation.[7] They suggest that the converging lines in the illusory figure correspond to converging lines which are typically and frequently produced in everyday three-dimensional viewing situations by corridors, tracks, roads, and so forth, and are normally associated with the observation of objects at a distance. To maintain size-constancy, a 'correction' is assumed to be made in the presence of converging lines with the result that the apparent size of an object will increase with its distance from the *S*. In three-dimensional situations, this correction compensates for the diminuton in retinal-image size with distance and, as such, produces size-constancy. Although this mechanism is of obvious utility in relation to size-constancy in everyday viewing situations, it is assumed to be inappropriately applied when viewing the Ponzo figure.

---

[6] Leibowitz and Heisel, *op. cit.*, 328–331; Hanley and Zerbolio, *op. cit.*, 437–452.

[7] R. Tausch, Optische Täuschungen als artifizielle Effeckte der Gestaltungsprozesse von Grössen-und Formenkonstanz in der naturlichen Raumwahrnehmung, *Psychol. Forschung*, 24–25, 1954, 299–348; E. von Holst, Aktive Leistungen der menschlichen Gesichtswahrnehmung, *Stud. Gen.*, *10*, 1957; H. L. Teuber, Perception, in *Handbook of Physiology*, 3, 1960, 1654–1657; Leibowitz, *Visual Perception*, 1965, 42–49; R. L. Gregory, *Eye and Brain*, 1966, 145–156.

The data from childhood through adulthood are consistent with this hypothesis. Although comprehensive developmental data on size-constancy are not available for populations comparable to those tested in the present study, experiments with children and adults as *S*s suggest that size-constancy increases with age and levels off in the early 'teens, after which it remains stable.[8] The similarity of the developmental functions for the Ponzo illusion and size-constancy is consistent with the point of view that they are both mediated by a common mechanism to some extent.

The decrease in the illusion with old age does not, however, fit readily into this theoretical framework. Although systematic data on size-constancy in old age are not available, the data do not suggest marked differences in size-constancy among adults of different ages.[9] From a theoretical point of view, it would be unreasonable to expect that size-constancy for distant objects, inasmuch as it is assumed to result from experience, would decrease with advancing age in the absence of a specific visual defect. It should be pointed out that size-constancy probably is mediated by a number of mechanisms,[10] and that the identification of a single factor common to the Ponzo illusion and size-constancy, *e.g.* perspective, does not necessarily imply that both phenomena are subserved by one and the same process throughout the entire life-span. The present data suggest only that these two phenomena share *at least* one common mechanism. Clearly, developmental functions for both measures over a wide range of ages and experiments are needed.

Summary. The magnitude of the Ponzo illusion was investigated for *S*s ranging in age from 3.5–88 yr. The illusion increases rapidly up to about age 13 yr., after which it remains stable through age 50 yr. and then decreases markedly. The possible relation between the Ponzo illusion and size-constancy for distant objects is discussed.

---

[8] H. P. Zeigler and H. W. Leibowitz, Apparent visual size as a function of distance for children and adults, this JOURNAL, 70, 1957, 106–109.

[9] Leo Postman and Edward Tolman, Brunswik's probabilistic functionalism, in *Psychology: A Study of a Science*, Sigmund Koch (ed.), I, 1959, 539.

[10] H. W. Leibowitz and D. Moore, Role of changes in accommodation and convergence in the perception of size, *J. opt. Soc. Amer.*, 56, 1960, 1120–1123.

INTERTRIAL INTERVAL AND RESPONSE-COMPETITION IN DISCRETE-TRIALS LEVER-PRESSING

By JOHN R. PLATT and KENNETH W. SPENCE, University of Texas

It has been realized for some time that a striking number of differences exist in the behavioral laws resulting from free-operant Skinner boxes and instrumental runways. The Skinner box does, however, offer a number of advantages, such as freedom from experimenter-produced variability, easily manipulable and physically-homogeneous response chains, extensive control and manipulative potential of both the inter- and intra-trial stimulus-situation and high efficiency due to its conduciveness to automatic data acquisition and analysis.[1] Thus, several recent studies have utilized discrete-trials lever-pressing (*DTLP*) situations in attempts to obtain results comparable to those from runways without sacrificing the numerous advantages of the Skinner box.[2] The results of these studies offer sufficient encouragement to justify further investigation and development of such techniques.

Two of the most obvious features which differentiate *DTLP* from runway situations are *S*'s presence in the situation during the intertrial interval (*ITI*) and the possibility of behavior involving the goal-cup's interfering with pressing. The first of these features suggests that intertrial variables such as *ITI* may play a crucial role in *DTLP*. The present study, then, investigated the effect of *ITI* on *DTLP* acquisition-performance while obtaining measures of competing behaviors involving the goal-cup.

Method. The *Ss* were 12 male and 12 female hooded rats approximately 6-mo. old, that had previously been used in an experiment on brightness-discrimination

---

* Received for publication October 11, 1966. This investigation was supported by a Public Health Service Predoctoral Fellowship (MH-28, 814–02) from the National Institute of Mental Health.

[1] For early discussions of the merits of such situations, see S. W. Bijou, The development of a conditioning methodology for studying experimental neurosis in the rat, *J. comp. physiol. Psychol.*, 44, 1942, 91–106; H. H. Kendler, Drive interaction: II. Experimental analysis of role of drive in learning theory, *J. exp. Psychol.*, 35, 1945, 188–198; J. H. Rohrer, Experimental extinction as a function of the distribution of extinction trials and response strength, *J. exp. Psychol.*, 37, 1947, 473–493.

[2] R. C. Gonzalez, Paul Bainbridge, and M. E. Bitterman, Discrete-trials lever pressing in the rat as a function of pattern of reinforcement, effortfulness of response, and amount of reward, *J. comp. physiol, Psychol.*, 61, 1966, 110–122; J. J. Porter and J. J. Hug, Effects of number of percentage of rewarded trials on the acquisition and extinction of lever pressing using a discrete-trial procedure, *J. exp. Psychol.*, 70, 1965, 575–579.

in the runway. The *Ss* were maintained throughout the present experiment on a ration of 10 gm. per day which had already been in force for 3 mo.

The apparatus was a box 9 × 9 × 12 in., with a grid-floor, opaque sides, and a clear Plexiglas top. A 1 × ¾ × ½-in. brass goal-cup, mounted in the center of the response-panel at grid-level, was the anode of a contact-sensing circuit through *S* to the grid. A Foringer rotary-solenoid retractable lever was mounted 3 in. above the floor and 2½ in. to the right of the goal-cup. The lever, which extended ¾ in. into the chamber, could be operated with a force of 15 gm. The box was placed in a refrigerator-hull along with a Foringer pellet-dispenser. The situation was programmed by a Teletype punched-tape reader and relay-circuitry. Data were acquired on a time-based, 3-channel printing counter which, along with the programming equipment, was located in a room adjacent to that containing the refrigerator-hull.

On the first day, all *Ss* were given 50 pellets in a free-operant, demonstration-type Skinner box. Each press delivered one 45-mg. Noyes pellet and a pellet was also dispensed whenever a press failed to occur within 60 sec. of the last pellet. The *Ss* of each sex were then assigned randomly and equally to 4 *ITI* conditions: 1, 5, 10, and 30 sec. The *ITI* was timed from the retraction of the bar to its reëntry. For the remainder of the experiment, each *S* was run every third day.

On Day 2, each *S* was given 100 continuously reinforced trials in the experimental chamber with the appropriate *ITI* in force. The trial was terminated and a pellet delivered whenever a press failed to occur within 60 sec. after the lever entered the chamber. Days 3–8 were identical to Day 2 except that only 50 trials per day were given and the 60-sec. criterion was eliminated. Throughout the experiment, presses made while *S* was in contact with the goal-cup were not reinforced, nor did they terminate the trial. This restriction was adopted to eliminate any ambiguity in respect to the incompatibility of goal-cup contact and lever-pressing. Measures were taken of the time between the introduction of the lever and *S*'s response to it, the total amount of time on each trial that *S* spent in contact with the goal-cup while the lever was in the chamber, and the total number of presses on each day which were made while *S* was in contact with the goal-cup.

Results. One *S* from Group 30 became ill on Day 7, but its previous data were retained. Figs. 1–2 present the mean learning curves for the speed of the responses and the reciprocal duration of the competing goal-cup contact-durations. These curves represent daily group-means of speed-transformations of the median times for individual *Ss* in 5-trial blocks. The different scales for the two measures reflect the fact that response-latencies were on the order of 10 times as long as competing goal-cup contact-times.

Analysis of variance of the response speed data yielded significant effects of Groups ($F = 4.97$, $df = 3/16$, $p < 0.025$), Days ($F = 65.92$, $df = 5/78$, $p < 0.001$), Groups X Days ($F = 2.14$, $df = 15/78$, $p < 0.025$) and Groups X Days X Sex ($F = 4.17$, $df = 15/78$, $p < 0.025$). A multiple-range analysis indicated that all the differences between the groups were significant well beyond the 0.05 level except for that between Groups 5 and 30. Thus, for the values employed in this study, response-speed was found to be a nonmonotonic function of *ITI* with a maximum at 10 sec. and a minimum at 1 sec.

Similar analyses of the competing goal-cup-contact speeds demonstrated significant effects of Groups ($F = 327.25$, $df = 3/16$, $p < 0.001$), Days ($F = 44.65$, $df = 5/78$, $p < 0.001$) and Groups X Days ($F = 2.69$, $df = 15/78$, $p < 0.005$). Differences between the group means were all significant well beyond the 0.05 level except for that between Groups 5 and 10. Thus, amount of response-competition from goal-cup-contact was a non-monotonic function of *ITI* with a minimum between 5 and 10 sec. and a maximum at 1 sec.

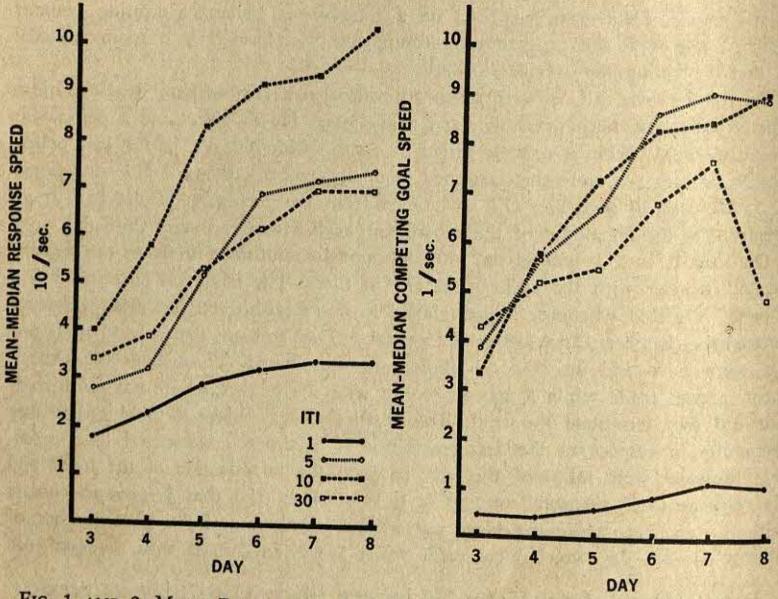


Fig. 1 and 2. Mean Reciprocal Median Response Speeds and Competing Goal-Cup Contacts
(Sessions 3 through 8)

The numbers of presses (nonreinforced) made while *S*s were in contact with the goal-cup decreased sharply from a mean of 50.1 on Day 3 to 8.2 on Day 5 and 1.6 on Day 8. The group means for this measure did not differ from each other after the first session.

Discussion. In general, the data suggest a negative correlation between response speed and the time of contact with the goal-cup. That so small a proportion of the latent period is involved in contact with the goal-cup is perhaps attributable to two facts: (a) a more realistic measure of the extent to which behaviors involving the goal-cup accounted for the variability in response-speed would have to include time spent traversing the distance between the cup and the lever, and also time spent hovering over the cup without actually contacting it; and (b) the response-differentiation provided by nonreinforcing presses made while in contact with the cup obviously produced some inhibition of intratrial goal-cup-contacting.

A closer examination of Fig. 1 suggests that the extreme slowness of Group 1 in respect both to pressing and to goal-competition is attributable to its *ITI* being shorter than its consumatory responses. It is to be noted that this group most closely approximates a free-operant condition and indicates the potential confounding involved in the use of free-operant Skinner boxes to study the relationship between rate of pressing and any variable affecting consumatory time such as the number of pellets presented as reinforcement.

The slower pressing speeds of Group 5 as compared with Group 10 can not be attributed to differential goal-competition, since no difference was found between these groups on the latter measure. Some refractory-period mechanism in the form of transient satiation, fatigue or the like might be called upon for explanation.

The non-monotonicity of the function relating pressing speed to *ITI* as a result of Group 30's depression appears to be rather complicated. First, note that this depression is correlated with an increase in goal-cup-competition. It was also observed during the study that this Group was less likely to be oriented to the lever at the onset of a trial than were Groups 5 and 10. This result suggests the need for a warning signal as *ITI*s become longer than several seconds. Further, it appears that as Group 30 learned to anticipate the end of the *ITI*, miscellaneous competing responses at the beginning of a trial came to be displaced by orientation to the goal-cup rather than to the lever. This aspect of the results is shown most clearly in Fig. 2 by the drastic increase in goal competition on Day 8 without any corresponding decrement in pressing speed.

The goal-competition exhibited by Group 30 cannot be attributed to actual consumatory behavior and must be listed as a fractional-anticipatory version of such behavior. Since most variables (*e.g.* time of deprivation or quantity or quality of reward) which would be expected to strengthen lever pressing would also tend to increase the tendency to approach and contact the goal-cup, several problems arise. When any such variable is being manipulated, its effects might well be underestimated due to confounding with competing anticipatory goal-behavior. Furthermore, within the framework of a theory such as put forward by Spence,[3] if the variables being manipulated were of a motivational nature (acting multiplicatively on associative factors) and the goal-approach and goal-contact habits were slightly stronger than the pressing habit, an actual inversion of a function relating lever-pressing to the independent variable might be expected. Such effects may well explain such unexpected free-operant results as the non-monotonic relationship found between time of deprivation and pressing rate by Belange and Feldman[4] and the elation effects found by Collier and Marx with incentive magnitude shifts.[5]

The problems indicated in this paper with respect to competition between lever-pressing and behavior involving the goal-cup lead to the conclusion that free-operant results may not be made comparable to those of the runway simply by the use of a retractable level producing discrete-trials. In view of this fact, we have devised a method in which a goal-container is located behind the lever-panel. Access to this container is through a circular opening just above the lever and large enough for *S*'s head to pass through. This opening is covered from behind the panel by a guillotine-type door operated by a synchronous motor. Preliminary data indicate that this spatial contiguity between lever and goal-container combined with controlled access to the container is effective in the removal of competition between pressing and goal-cup behavior.

SUMMARY

Four groups of rats were trained in a discrete-trial lever-pressing apparatus with intertrial intervals (*ITI*) of 1, 5, 10, or 30 sec. Measures of goal-cup contact-time and of response-time varied significantly with *ITI* in a manner which reflected the competition of lever-pressing and goal-cup behavior. Several methodological suggestions were made on the basis of these results.

---

[3] K. W. Spence, *Behavior Theory and Conditioning*, 1956, 93, 133–137.

[4] David Belanger and S. M. Feldman, Effects of water deprivation upon heart rate and instrumental activity in the rat, *J. comp. physiol. Psychol.*, 55, 1962, 220–225.

[5] George Collier and M. H. Marx, Changes in performance as a function of shifts in the magnitude of reinforcement, *J. exp. Psychol.*, 57, 1959, 305–309.

EFFECTS OF THE RELEVANCE OF INTERPOLATED STIMULI ON TIME-ERROR IN THE COMPARATIVE JUDGMENTS OF WEIGHTS

By WILLIAM E. GUMENIK and BRUCE WEISS, The University of Toledo

Negative time-errors commonly occur in psychophysical judgments of weights.[1] If, however, a weight heavier than the standard is interpolated between the standard and the comparison-stimulus, the Point of Subjective Equality (*PSE*) is shifted upward, sometimes to the degree that positive time-errors occur.[2] This upward shift is relatively small, indicating that the extraneous weight contributes less than the standard to the *PSE*. Guilford and Park, using a 200-gm. weight as the standard, found that the interpolation of a 400-gm. weight resulted in an upward shift in the *PSE* of only 7 gm.[3]

The upward shift in the *PSE* is explainable by the theory of adaptation-level (*AL*), which holds that all judgments are made in reference to the *AL*, rather than to any specific stimulus.[4] All present and past stimuli, including the interpolated weight as well as the standard, determine the *AL* and therefore the *PSE* (with which it theoretically coincides). The unequal contribution of the standard and the interpolated weight to the *PSE* can be subsumed under the theory's differential weighting of focal, background, and residual stimuli.

Gleitman and Hay hypothesized that the unequal contribution of the standard and the interpolated weight to the *PSE* might be attributable to greater attention toward the standard, since only the standard was relevant to the judgment task.[5] This hypothesis was based in part on an earlier finding by Brown that weights which were judged, and therefore focal in attention, affected series judgments more than weights which were not judged.[6] Gleitman and Hay, by not revealing to *S*s which of two weights was the standard until both were presented (post-instruction condition) forced the *S*s to attend to both. Contrary to their predictions, forced attention to the heavier extraneous weight resulted in lower *PSE*s than the pre-instruction condition, in which the *S*s were instructed as to which weight was the standard prior to presentation of the weights.

---

* Received for publication September 2, 1966.

[1] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 225–227.

[2] J. P. Guilford and D. G. Park, The effect of interpolated weights upon comparative judgments, this JOURNAL, 43, 1931, 589–599; Henry Gleitman and J. C. Hay, The effect of instruction on positive time-errors in successive comparison, this JOURNAL, 77, 1964, 286–290.

[3] Guilford and Park, *op. cit.*, 592.

[4] Harry Helson, *Adaptation-Level Theory*, 1964.

[5] Gleitman and Hay, *op. cit.*, 286–290.

[6] D. R. Brown, Stimulus-similarity and the anchoring of subjective scales, this JOURNAL, 1953, 66, 199–214.

Gleitman and Hay's *ad hoc* explanation of their results maintained that while the post-instruction conditions increased attention to the interpolated weight, this attention was directed toward the explicit non-relevance of that weight for the series. This reduced the contribution of the interpolated weight to the *PSE*. They reconciled their results with those of Brown by holding that "Brown's categories of 'judged' and 'not judged' stimuli correspond to the differences between positive relevance and neutral status";[7] while in their study, "no judged" (the effect of pre-instruction) and "judged separately" (the effect of post-instruction) "may correspond instead to the difference between neutral status and explicit non-relevance."[8] If, however, Gleitman and Hay's reasoning is carried to its logical conclusion, Brown's findings also might be completely attributable to the explicit non-relevance of his non-judged weights, since in Brown's experiment the *S*s were specifically told "don't judge" each time a non-judged weight was presented.[9]

The Gleitman and Hay experiment also did not contain a control in which no extraneous weight was presented, hence it could not be determined whether explicit non-relevance completely eliminated the effects of the interpolated weights.

The present experiment was designed to investigate the effects of increased attention to a heavy interpolated weight as a relevant member of the stimulus-series, as well as the effects of explicit non-relevance of the interpolated weight on the *PSE*. It is thought that forced judgment of interpolated weights, under conditions in which the same comparison-stimulus is used for the judgment of both the standard and the interpolated weight sequentially, should heighten attention toward the interpolated weight as a relevant member of the stimulus-series. On the other hand, explicitly telling *S* "don't judge it" before each presentation of the interpolated weight should render this weight explicitly non-relevant for the stimulus-series. The *PSE*s obtained under these conditions of relevance and explicit non-relevance of the interpolated weight were compared with each other and with the *PSE*s obtained under the usual testing condition, in which no judgment of the interpolated weight is called for, but in which the *S*s were not explicitly told "don't judge it" each time the interpolated weight is presented. The *PSE*s were also compared to a control condition in which no interpolated weights are presented.

Method: (1) Subjects. The *S*s were 80 undergraduate students, randomly assigned to four groups of 20 *S*s each.

(2) Materials. The stimuli were cylindrical Kodak film canisters, approximately ¾ in. in height and 2 in. in diameter, filled with buckshot and cotton. The standard was 100 gm.; the interpolated weight, 200 gm., and the 5 comparison-weights, 80, 90, 100, 110, and 120 gm.

Procedure. Group 1 (Relevant). The *S*s individually received 50 test-trials, in which the standard was paired 10 times with each of the 5 comparison-weights. The order of the comparison-weights was so randomized in blocks of 5, that each of the 5 different comparison-weights was paired with the standard before the next randomized block began. The presentation of the standard preceded the presentation of the comparison-weight by 8 sec., and the intertrial interval was 12 sec. The interpolated weight was presented between the presentation of the standard and comparison weights, 4 sec. after the standard and 4 sec. before the comparison-weight. The *S*s were required to lift each weight and immediately put it down without hefting it. After presentation of the comparison-weight, it was judged against the standard and interpolated weights in the order of their presentation on the following 9-point scale: very much lighter, much lighter, lighter, a little lighter, equal, a little heavier, heavier, much heavier, very much heavier. These categories were printed on a card in an easily readable position in front of *S*. The 50 test-trials were preceded by 3 practice trials, one each with the 80-, 100-, and 120-gm. comparison-weights.

---

[7] Gleitman and Hay, *op. cit.,* 286–290.
[8] *Ibid.,* 290.
[9] Brown, *op. cit.,* 203.

Group 2 (Neutral). The procedure with Group 2 was the same as with Group 1 except that while *S* lifted the interpolated weight as well as the standard he was only instructed to judge the third weight (comparison-weight) against the first weight (standard). No judgment of the interpolated weight was required.

Group 3 (Non-relevant). The procedure with Group 3 was the same as with Group 2 except that *S* was told "don't judge it" each time the interpolated weight was presented.

Group 4 (Control). The procedure with Group 4 was the same as that of the preceding groups except that no interpolated weight was presented.

The *PSE*s based on the judgments of the comparison-weights against the standard were computed for each *S* in each group by the average z score method for the method of constant stimuli,[10] with any of the four "lighter than" categories being considered a "lighter than" response, any of the four "heavier than" categories a "heavier than" response and the "equal" category an "equal" response.

Results. The mean *PSE*s for the relevant, neutral, non-relevant and control groups respectively were 102.3, 102.0, 97.7, and 97.5 gm. A randomized group-analysis of variance revealed this effect to be significant ($F = 3.69$, $df = 3$ and 76, $p < 0.05$). Individual comparisons among groups were made by Duncan's new multiple range-test.[11] The relevant and neutral groups had significantly higher mean *PSE*s than the non-relevant and control groups at the 0.05 level. The differences between the relevant and neutral groups and between the non-relevant and control groups were not significant at the 0.05 level.

Discussion. The relevant and neutral conditions yielded significantly higher *PSE*s than the control and non-relevant conditions and did not significantly differ from each other. It therefore appears that a weight, heavier than the standard, which is interpolated between the standard and comparison-weights and presented under conventional (neutral) conditions raises the *PSE* of the standard. Forced judgment of the interpolated weight does not increase or lessen this effect. Either interpolated weights are fully attended to as relevant members of the stimulus-series, whether or not *S* is required to judge them overtly, or increased attention to the interpolated weights as relevant members of the stimulus-series does not affect the *PSE*. In either case, the unequal contribution of the standard and the extraneous interpolated weight to *PSE* can not be attributed to greater attention to the standard.

---

[10] Woodworth and Schlosberg, *op. cit.*, 205–206, 214–215.

[11] A. L. Edwards, *Experimental Design in Psychological Research*, 1962, 136–140.

The fact that the non-relevant condition yielded significantly lower *PSEs* than the neutral condition supports Gleitman and Hay's contention that explicit non-relevance of the interpolated weight reduces the effect of this weight on the *PSE* of the standard.[12] In fact it appears that explicit non-relevance completely removes the effect of the interpolated weight since the *PSEs* obtained under the non-relevant condition did not differ from those obtained under the control condition.

The finding of no differences between judged (relevant) and non-judged (neutral) interpolated weights on the *PSE* appears to be in opposition to the results of Brown, who found that judged weights had significantly greater effects on the *AL* (which theoretically coincides with the *PSE*) than non-judged weights.[13] It is possible that Brown's findings might be completely attributable to the explicit non-relevance of the non-judged weights, rather than to increased attention to the judged weights, since Brown's *Ss* were explicitly told "don't judge" each time a non-judged weight was presented. There are, however, many differences between the present procedure and Brown's, including Brown's use of the method of single stimuli. It is possible that increased attention to extraneous stimuli might have greater effects when no standard is presented.

SUMMARY

This experiment investigated the effects, on the *PSE*, of interpolating a weight, heavier than the standard, between the standard and comparison weight, under three different degrees of relevance of the interpolated weight. In the relevant condition, attention to the interpolated weight as a relevant member of the stimulus-series was increased by forcing *S* to judge the same comparison weight against the standard and interpolated weights sequentially. In the neutral condition, *S* was not required to judge the interpolated weight overtly. In the non-relevant condition, the interpolated weight was rendered explicitly non-relevant by *E*'s stating "don't judge it" each time the interpolated weight was presented. These conditions were compared with a control condition in which no interpolated weight was presented. Presentation of the interpolated weight raised the *PSE* under the neutral condition, and forced judgment of the interpolated weight under the relevant condition did not change the extent of this effect. Explicit non-relevance of the interpolated weight completely removed the effect of the interpolated weight on the *PSE*.

---

[12] Gleitman and Hay, *op. cit.*, 290.
[13] Brown, *op. cit.*, 203.

APPARATUS

GROUP EXPERIMENTS IN HEARING

By ROBERT J. GRISSOM, San Francisco State College

Judging from the number of laboratory manuals which include exercises in hearing, there is considerable interest in the area in general and experimental psychology courses.[1] Unfortunately, where the apparatus budget is limited such exercises are restricted to rather casual demonstrations of elementary phenomena. Highly versatile and precise apparatus would likely be out of reach of many courses. This note presents one possible compromise between cost and rigor, with some versatility added.

Typical equipment for group demonstrations includes the Seashore Musical Aptitude Records, clinical audiometer, and an audio oscillator-amplifier-speaker combination. Each of these choices has its limitations. Seashore record tones emanating from a speaker are subject to the speaker's operating characteristics, which are inferior to those of earphones, and to reflection from the walls and objects in the room. The consequent superposition and interference of waves results in a signal which is no longer a simple tone and which differs for each student according to his position in the room. The only advantage of such a demonstration is its relative inexpensiveness.

The use of clinical audiometers with earphones avoids shortcomings of loudspeaker presentations just considered. Clinical audiometers are designed, however, to determine *hearing loss* at a limited number of fixed frequencies. To obtain his absolute thresholds, the student must be given an average audibility curve and then subtract his hearing-loss readings from the normal thresholds at each frequency. Furthermore, stages of attenuation are a rather gross 5 decibels in the commonly used Maico audiometer and, since the purpose of the audiometer is clinical, the range of attenuation suits the purpose of detecting elevated thresholds. It is possible that students with sensitivity much above average may not be able to attenuate sufficiently to reach threshold at some frequencies. For example, persons with thresholds more than 10 db. below normal cannot attenuate to threshold using the Maico MA-12 audiometer. Assuming a laboratory class of 24 students, for cost comparisons with other systems, 12 Maico MA-12 audiometers would cost approximately $4,250.

---

[1] L. M. Baker, C. Weisiger, and W. M. Taylor, *Laboratory Experiments in General Psychology*, 1960 21–28; Harlan Lane and Daryl Ben, *A Laboratory Manual for the Control and Analysis of Behavior*, 1965, 83–157; J. H. L. Roach, B. R. Sher- man, and P. P. Roach, *Experiments in General Psychology*, 1965, 51–61; J. N. Shafer, *Laboratory Exercises in Psychology* 1965, 154–158; J. C. Stevens, R. J. Herrnstein, and G. S. Reynolds, *Laboratory Experiments in Psychology*, 1965, 23–36.

Recent years have witnessed the development of direct psychophysical methods, many of which have earned their inclusion in laboratory manuals.[2] The apparatus suggested by the authors of these manuals consists of an audio oscillator, attenuator(s), vacuum tube voltmeter (*VTVM*), and earphones for each pair of students. This apparatus is suited not only for psychophysical scaling, using the direct methods, but for classical threshold, generalization, discrimination, fatigue, and audiogram determinations. Using the least expensive components commercially available, the cost of equipping 12 pairs of students would likely be $2,100, or about half of the cost of 12 audiometers. The high quality, but moderately priced, system used in the elementary laboratory at the University of Michigan consists of Hewlett-Packard audio frequency generators (Model 200-AB), Daven fixed and variable attenuators (Models T-691 and T-349-G), Simpson Electric VTVMs (Model 715 AC), and Grason Stadler earphones (Model TDH-39).[3] The cost for 12 sets of this equipment is approximately $4,100. Thus, for less than the cost of audiometers, an auditory system can be connected which provides continuous variation of frequency across the range of hearing, accurately monitored variation of intensity in fine steps to absolute threshold, and calibrated earphones.

An alternative system has been built for the Department of Psychology at San Francisco State College.[4] The arrangement consists of a master control booth and listening station and six two-subject listening stations. Each of the stations is a modified language laboratory booth with acoustical dividers (Thompson, Ramo, Wooldridge, booth model STB). The control booth contains two audio signal generators (Hewlett Packard, model 200-AB) and two audio amplifiers (McMartin, model MA 20). The station of each *S* consists of two earphone-jacks; dynamic stereo earphones (Superex, model ST-M); a lever switch for monaural, binaural, or dichotic presentation; a 10-step attenuator governing a range of 100 db.; and a fine attenuator ranging from —30 to 10 db. monitored by a zero-centered db. meter (Argonne VU meter, Model 331). A schematic diagram of this system is presented in Fig. 1 with a list of parts.

---

[2] Lane and Ben, *op. cit.*, 149–157; Stevens, Herrnstein, and Reynolds, *op. cit.*, 1965, 26–41.

[3] Lane and Ben, *op. cit.*, x-xii.

[4] This system was built by the Educational Electronics Division of McCune Sound Service, San Francisco.

The system just described is suitable for each of the experiments listed

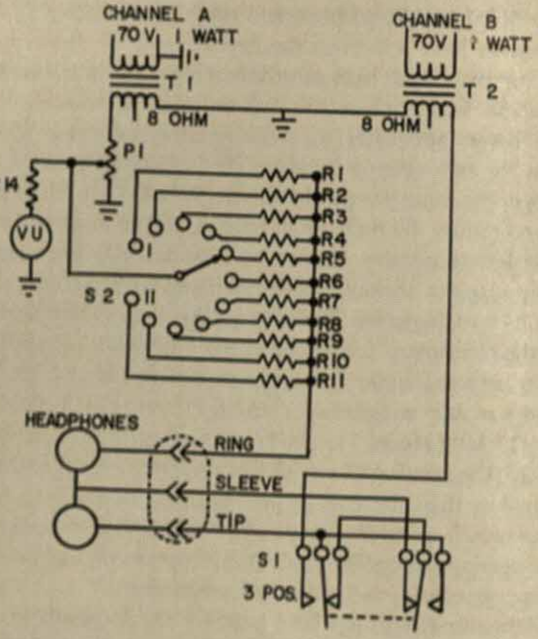


FIG. 1. WIRING DIAGRAM OF *S*'S LISTENING STATION

The two amplifiers are located at the master control booth. The attenuation at the master listening station ranges from R_1 through R_{13} to accommodate *Ss* with exceptionally low thresholds. There is no potentiometer (P_1) at the master listening station.

PARTS AND SPECIFICATIONS

P_1	1 KΩ	R_5	1.8 KΩ	R_{10}	560 KΩ	S_1 Switchcraft 1307
R_1	10 Ω	R_6	5.6 KΩ	R_{11}	1.8 MΩ	S_2 11 Pos NS
R_2	47 Ω	R_7	18 KΩ	R_{12}	5.6 MΩ	T_1 and T_2 Peerless 15064
R_3	180 Ω	R_8	56 KΩ	R_{13}	15 MΩ	
R_4	560 Ω	R_9	180 KΩ	R_{14}	3.6 KΩ (VU)	

above as possibilities with the previously described oscillator-attenuator-VTVM-earphone arrangement, and, since the present arrangement is two-channel, additional experiments are possible. A variety of dichotic and beats phenomena may be demonstrated. Also, the instructor or experimenter can use one channel for giving instructions, or tape recordings can be played through the system. For example, two groups of *Ss* can be given different taped instructions simultaneously when each group is presented with a different channel. A graduate student has used the apparatus for a group experiment in the semantic-differential scaling of taped tonal phrases.

The major uses to which the system has been put thus far are in undergraduate demonstrations of the U-shaped audibility curve and in the scaling of loudness using the method of magnitude-estimation.[5] In the latter experiment, a linear relation between log intensity and log loudness has been consistently demonstrated in individual and group data, confirming Stevens' Power Law. Slopes for group data have ranged from 0.4 to 0.7.

In comparison with the oscillator-attenuator-VTVM-earphone arrangement, the present system involves the added expense of language laboratory booths and two amplifiers. These costs are, however, more than offset by the use of only two oscillators rather than an oscillator for each student.

The present system was built primarily for undergraduate laboratory experiments in classes of about 24 students working in pairs, and was limited to an expenditure of approximately $2,900. This apparatus is thus considerably less expensive than 12 audiometers or 12 oscillator-attenuator-VTVM-earphone combinations. Where more money is available, the present design could be built with more reliable research-quality components.

Each attenuator resistor had its resistance confirmed by an ohmeter, and representative frequencies were monitored for sinusoidality across intensities by a Tektronix oscilloscope (Model 531 A). Visual inspection of the waveform confirmed auditory tests with the earphones to the effect that good sine waves were being generated.

For research purposes, each level of attenuation at each station should be checked to a strict tolerance on a high-quality oscilloscope. On this test, the present system performed adequately for its purposes (*i.e.* each step of the attenuator effected approximately a 10-fold change in intensity). An additional refinement for research purposes would be substitution for its circumaural cushion earphones by individually calibrated sponge rubber cushion earphones (*e.g.* Telephonics, model TDH-39 or TDH-49). An average calibration curve is available from the manufacturer of the Superex ST-M earphones used in the present system.

---

[5] Stevens, *et al.*, *op. cit.*

NOTES AND DISCUSSIONS

THE MEASUREMENT OF SUPRATHRESHOLD PAIN

The study of both subjective and physiological responses to painful stimulation is fraught with a number of difficulties, one of which is the difference in functional relationships between the threshold and suprathreshold of pain, such as surgical pain. For example, morphine relieves surgical pain in some instances without having an effect upon the threshold of pain.[1]

Before undertaking studies of hypnotic analgesia we thought it desirable to calibrate the suprathreshold of pain in the normal waking state. This note refers to some preliminary results obtained in the course of the study when the *S*'s hand and arm are immersed in ice water. In a study of pain aroused by a tourniquet, Beecher and others found that they could make use of a subjective scale extending from '0' to '4' with '4' being unbearable.[2] According to a procedure which they standardized, following finger exercise while the arm circulation was stopped, pain mounted gradually according to the subjective reports, and usually became unbearable, although the time to reach maximal intensity varied from 3–47 min. with different *Ss*. We have followed their lead in asking our *Ss* to report the increasing pain of the arm immersed in ice water as a function of time in the water, and have found that a lawful function results. It is our hope that we may study deviations from this function under hypnosis as a quantitative approach to the study of analgesia, and we have a few preliminary results to indicate that the method holds promise. This Note, however, is concerned with the non-hypnotic data.

Seventeen university students (9 men and 8 women) served as *Ss* in the experiment. After an explanation of the reason for undertaking the study in a laboratory devoted to hypnosis, the *S* was instructed to report pain on a scale of '0' to '10,' as follows:

Please report '10' when you can no longer keep your hand in the ice water. State '0' should be reported before the hand is immersed in the water. Your hand is at '0' when it is out of the water. When you place your hand in the water you will find that the arm gradually becomes more and more painful. Please report by stating a number from '0' to '10' as the pain increases. Report '10' when you can no longer keep your arm in the water.

---

* This study was conducted at the Laboratory of Hypnosis Research, Department of Psychology, Stanford University, supported by the National Institute of Mental Health (Grant MH-3859) and the Air Force Office of Scientific Research (Contract AF 49(638)-1436) under the direction of Ernest R. Hilgard.

[1] H. K. Beecher, *Measurement of Subjective Responses*, 1959, 111–130.

[2] Beecher, Pain: One mystery solved, *Science*, 152, 1966, 840–1. *Cf.* also K. M. Dallenbach, A pain-scale: A report from a victim of tic douloureux, this JOURNAL, 76, 1963, 152–155.

One might obviously have given specially motivating instructions, but these relatively matter-of-fact instructions were easily understood and accepted by the *Ss*.

The hand and forearm were immersed in a tank of water 215 × 460 cm., maintained at approximately 0°C. by having melting ice suspended in the water, with some salt added. It was necessary to stir the water repeatedly in order to prevent a layer of warmer water from surrounding the hand and arm.[3]

Of the 17 *Ss*, 8 of 9 men reached State 10 and removed their hands from the water in a period of time varying from 33–112 sec.; 7 of 8 women reached State 10 and removed their hands between 16–100 sec. (Table I). Sex-differences, if any, are not pronounced enough to yield significance in this table ($t = 1.23$, n.s.). The man and woman who failed to reach State 10 both kept their hands in the water until removed at the end of the trial (4 min.) with final reports of Pain-states 8 and 9, respectively.

The most striking result was the relatively uniform course of the response-curve. There was typically some delay before any pain was reported: at first the water feels very cold but not painful; then reports of pain increased fairly rapidly, but slowed down, as shown in Fig. 1, as an asymptote was approached. At this stimulus-temperature and with our instructions, there do not appear to be any reversals; that is, the pain increases without any numbness developing to dull it.

The curves, while differing in detail, both appear to belong to a common family of sigmoidal curves. Following an analysis made by Voevodsky of the neural effects over time of a step-input (in this case, stimulating the receptors in the arm by a sudden change from room temperature to that of ice water),[4] it was predicted that the function could be fitted approximately by a curve representing a second-order linear system, expressed in differential form by the expression:

$$\Delta T = T_1 - T_2 = a \frac{d^2R_p}{dt^2} + b \frac{dR_p}{dt} + cR_p. \qquad \text{Equation [1]}$$

Here ΔT is the temperature change between the room temperature (T_1) and the temperature of the water (T_2), R_p is the pain-response, t is time in seconds, and a, b, and c are constants for the particular individual in the given setting. Through steps that need not concern us at this time,

---

[3] This somewhat crude apparatus is being replaced with a more carefully designed one involving circulating refrigerated water maintained thermostatically at the desired temperature.

[4] John Voevodsky, A mathematical representation of interneuron behavior, Doctoral dissertation, Stanford University, 1965.

TABLE I

TIME (IN SEC.) AFTER IMMERSION OF ARM IN ICE WATER FOR SUCCESSIVE REPORTS OF PAIN-STATES

		Pain-States										
	Ss	0	1	2	3	4	5	6	7	8	9	10
Men	A	0	(2)*	(3)	(4)	(6)	8	9	16	19	27	33
	B	0	4	6	8	10	12	15	19	27	31	40
	C	0	1	3	4	6	10	14	17	21	31	41
	D	0	(2)	(5)	7	13	17	21	27	32	37	49
	E	0	1	2	4	6	12	15	18	21	39	65
	F	0	7	(10)	12	(16)	20	27	37	49	58	72
	G	0	6	8	12	16	19	23	30	34	44	94
	H	0	5	8	11	14	19	24	33	42	67	112
	I	0	(2)	4	10	14	18	22	33	73	—	—
Sum (Omitting I)		0	28	45	62	87	117	148	197	245	334	506
Mean ($N=8$)		0	3.5	5.6	7.8	10.9	14.6	18.5	24.6	30.6	41.8	63.2
Women	J	0	2	3	4	6	7	10	11	12	14	16
	K	0	1	3	6	7	8	9	11	14	18	27
	L	0	3	4	6	8	9	11	13	15	24	35
	M	0	4	6	8	10	11	13	15	17	23	36
	N	0	5	(6)	7	8	11	15	21	28	36	43
	O	0	15	21	24	27	30	34	40	45	51	61
	P	0	4	12	18	(21)	24	35	42	54	68	100
	Q	0	7	13	21	29	35	42	47	66	92	—
Sum (Omitting Q)		0	34	55	73	87	100	127	153	185	234	318
Mean ($N=7$)		0	4.9	7.9	10.4	12.4	14.3	18.1	21.9	26.4	33.4	45.4

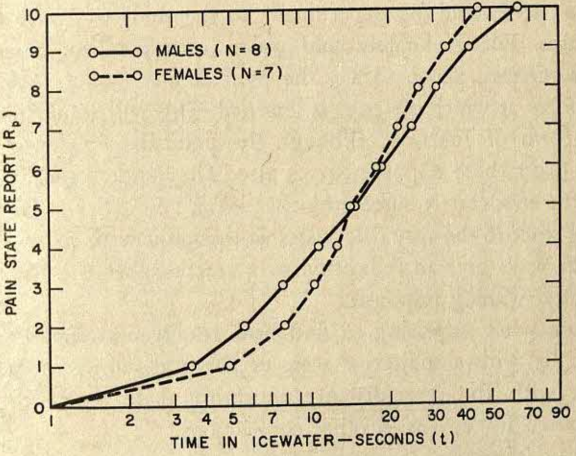
* Numbers in parentheses have been interpolated in order to permit averaging.

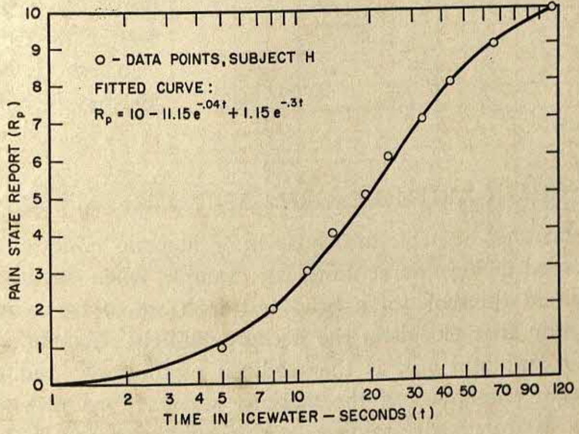
a transformation can be made so that the equation can be solved for R_p, the pain-response. The form of the resulting equation is:

$$R_p = A + Be^{-mt} + Ce^{-nt}. \qquad \text{Equation [2]}$$

Here A, B, C, m, and n, are constants.

It may be noted that the two exponential terms provide for double





inflections in the curves, found empirically in the somewhat sigmoidal form of the mean curves of Fig. 1.

As an illustration of the appropriateness of fitting a curve of this form to the data, the data from one S have been plotted in Fig. 2, along with the points fitted by the curve. The actual function fitted is as follows:

$$R_p = 10 - 11.15e^{-0.04t} + 1.15e^{-0.3t} \qquad \text{Equation [3]}$$

Because the data-points fall very near to the fitted curve, it appears that the selected function is an acceptable one.

When the improved equipment has been installed in the laboratory, we propose to extend the experiments to the study of water of different temperatures. For moderately cold water, corresponding, say, to that of the Pacific Ocean, about 15°C., the pain at first mounts and then there is a numbness at which the pain is lessened. This will result in a somewhat different form of function, although the particular equations chosen are predicted to fit these response-curves also. The question can then be raised whether the moderately susceptible hypnotized *S*, under analgesic instructions, will react to the very cold water as though it were somewhat warmer, or in other ways give an indication of a reaction that is lawful but differing from the waking response.

It is somewhat surprising to find how readily untrained *S*s can report systematically with a subjective scale of this sort, although the finding is consistent with the sense-distance reporting done across modalities by the *S*s studied by Stevens in other connections.[5]

Stanford University

JOHN VOEVODSKY
LESLIE M. COOPER
ARLENE H. MORGAN
ERNEST R. HILGARD

AN ANTILOGARITHMIC GRID AND SOME APPLICATIONS

It is sometimes desirable to display the arithmetic values of data originally recorded in logarithmic units, for example, when fluctuations in the visual masked threshold of a brief test flash are to be compared with the waveform that modulates the background-light. Comparisons of this kind are meaningful only if the readings of threshold and the related instantaneous amplitudes of the masking wave are made commensurate. If the raw data are wedge-readings at threshold expressed in the logarithmic units of optical density, the experimenter must ordinarily take the antilogarithms of the density readings, plot the resulting values on a linear ordinate, and finally adjust the scale of this ordinate to fit the range of the thresholds to the masking waveform as drawn. This note describes a simple plotting-grid (as in Fig. 1) that makes such maneuvers possible without tables or involved calculation and lends itself to a variety of other uses, some of which will be discussed.

---

[5] S. S. Stevens, The surprising simplicity of sensory metrics, *Amer. Psychologist*, 17, 1962, 29–39.

Theory and construction. An antilogarithmic axis is fundamentally a linear arithmetic scale divided into segments that represent successive steps of equal size on a logarithmic scale. (Note that logarithmic calibration is the converse, since each division of the logarithmic scale represents a step of some constant size along an arithmetic scale.) If each step is 1

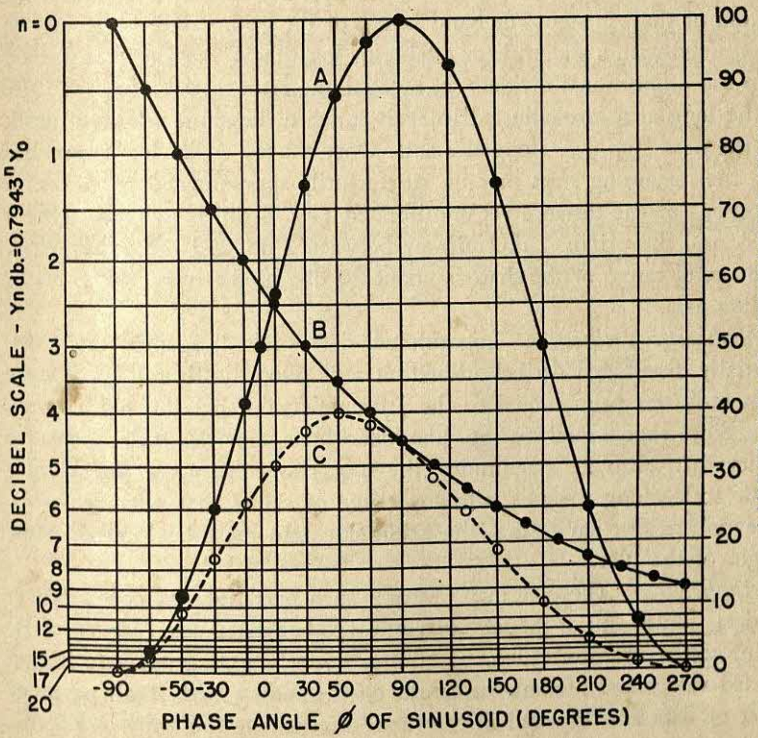


FIG. 1. TYPICAL ANTILOGARITHMIC GRID

Curve A illustrates the linearization of a set of logarithmic data. Plotted against the antilogarithmically distributed scale of decibels, values of $f(\phi) = 10 \log 0.5(1 + \sin \phi)$ assume ordinates proportional to $(1 + \sin \phi)$. Curve B is an exponential function that drops 0.5 db. for each advance of 20 deg. in ϕ. Curve C is the product of the ordinates of Curves A and B plotted according to the procedure for multiplication described in the text.

decibel (db.),[1] the upper and lower boundary-values of any given segment are in the ratio 1:0.7943. It follows that successive downward steps of 1 db. along an antilog axis diminish in size. On the y-axis, they are distributed according to the exponential equation

$$y_n = 0.7943^n y_o$$

where y_o is the over-all height of the axis, n the number of steps counted from the top of the axis, and y_n the distance between the origin and the bottom of the nth step. For hand layout of the grid, values of y_n may be determined with a slide rule or a table of logarithms. An alternate method is to use proportional dividers set by trial and error to the ratio 1:0.7943. If the long arms are adjusted to span y_n, then the short arms will mark off the next step y_{n+1} along the axis. Once the top 3 db. have been laid out, the remaining steps may be satisfactorily approximated by successive bisections of the distance yet uncalibrated (*e.g.* 4 db. is midway between 1 db. and the origin, 5 db. midway between 2 db. and the origin, etc.). The useful range of the chart is about 20 db., or the upper 99% of the antilog axis.

Arithmetic calibration. Logarithmic data plotted against the exponentially distributed decibel-calibration will automatically fall on a curve with ordinates proportional to the arithmetic values of the plotted variable. Since exponential spacing, like logarithmic, defines a scale of relative rather than absolute magnitudes, the user may construct a linear arithmetic scale along the axis with any range of values that suits his needs. The only restriction is that the arithmetic zero must be located at the origin of the plot.

Applications. The user may perform such operations as normalization of data, adjustment of the ratio of vertical to horizontal size (aspect-ratio), or projective transformation, by adding decibels appropriately to his recorded values and plotting the results on the antilog grid. If all points of a set of data are displaced in the same direction by the same number of decibels, their distribution on the antilog plot will be expanded or compressed but will retain its original internal proportions. The following procedures (1-3) are based on this fact.

(1) Normalization. First select the datum-point for which a standard position on the chart is desired, and add to its original value the number of decibels that will bring it to this position when replotted. Replot the remaining points after adding the same number of decibels to each. In this way two sets of data may be equated at any selected point and their conformity examined. Similarly, a plotted

---

[1] S. S. Stevens, Decibels of light and sound, *Physics Today,* 8, 1955, 12–17.

curve may be matched in size to a criterial curve. If the lowest arithmetic value of both curves is zero, merely raise or lower each of the data-points by the difference in decibels between the maxima of the two curves as read from the antilog scale. If either minimum differs from zero, replot so as to bring it to zero, *i.e.* by uniform vertical translation of the entire curve, and then add decibels as indicated. For this purpose it is convenient to have the antilog grid movable relative to the data, on a tracing table or as a transparent overlay.

(2) Adjustment of aspect-ratio. The procedure is the same as that for vertical equalization of two curves. After displacement of the data to locate the minimal value at arithmetic zero, the user reads the difference in decibels between the actual and desired positions of the maximum and then raises each displaced datum-point by this number of units along the antilog axis. In this example we assume that the length of the curve along the abscissa is fixed.

(3) Projective transformation. A geometric figure is distorted if projected upon a plane inclined to its original plane. For example, a slanted cross-section of a right circular cylinder has a height greater than the diameter of the cylinder; conversely, the image of a circle on the ground glass of an obliquely-aimed camera is foreshortened. The antilog grid enables the user to plot with little effort figures that approximate the results of such projections. For simplicity let the initial figure be in a vertical plane, the line of projection be normal to the figure (think of the axis of the right circular cylinder), and the plane onto which the figure is projected be tilted about an axis parallel to the abscissa to form an angle Θ with the initial plane. The height h' of the figure as projected will be related to the original height by the expression

TABLE I

RATIOS OF PROJECTED TO INITIAL HEIGHT AND THEIR EQUIVALENTS IN DECIBELS

A figure of height h has a projected height $h' = h/\cos\theta$ upon a plane inclined θ degrees to its original plane.

θ	h'/h	Decibels	θ	h'/h	Decibels
0	1.000	0.00	55	1.744	2.41
10	1.015	0.07	60	2.000	3.01
20	1.064	0.27	65	2.37	3.74
30	1.155	0.62	70	2.92	4.66
40	1.306	1.16	75	3.86	5.87
45	1.415	1.50	80	5.78	7.60
50	1.555	1.92			

$$h' = h/\cos\Theta$$

Table I gives the ratio h'/h in decibel form for typical values of Θ. For projections with the properties of slanted cross-sections, raise each corner or other informative point of the figure on the plot by the number of decibels, as nearly as possible, indicated for the desired angle.[2] For projections that share the properties of the image in a camera the geometry is inverted; the initial plane is in effect slanted, the plane of projection vertical, but the line of projection remains normal to the vertical plane, and points on the figure should be lowered on the chart.

---

[2] Fred Attneave, Some informational aspects of visual perception, *Psychol. Rev.*, 61, 1954, 183–193.

(4) Multiplication and division. The exponential curve and the damped sinusoid in Fig. 1 illustrate the use of the antilog grid to display products and quotients of functions. The damped sine wave C is the sinusoid A multiplied by the exponential B, and might represent the output of a sinusoidally-excited system of transmission whose gain decreases exponentially during the period of observation. The operations are those of logarithmic addition or subtraction. To multiply any two functions plotted on the grid, read from the antilog axis the ordinates in decibels of points on both functions with the same abscissa, and plot their sum. Repeat for as many pairs of points as are needed to delineate the functions adequately. To divide, subtract the values read for points on the divisor function from those of corresponding points on the dividend function and plot the differences. If any such difference is negative, it will be necessary to shift the zero of the decibel scale downward sufficiently to confine the plotted points within the chart. It may be seen that although the result of multiplication or division will have the form of the product or the quotient in linear coördinates, it will probably have to be expanded or reduced or given a different arithmetic scale in order to occupy the correct range of numerical values.

The antilog x-axis. The foregoing examples suggest the utility of the antilog *y*-axis for such purposes as the display or manipulation of data or the construction of stimulus-figures or illustrations. Additional possibilities emerge from an antilog calibration of the *x*-axis. Some typical uses include normalization of psychometric functions with different ranges or derived from different modalities; comparisons between waveforms that differ in period; and construction of normalized learning curves (such as Vincent curves).[3] With both axes calibrated as antilog scales, the grid may serve as a pantograph for point-by-point enlargement or reduction of figures or as a 'viewer' for projective transformations that involve planes at compound angles. Two-way antilog calibration also enables the user to generate families of equal-area figures or to multiply the area of a given figure by a constant. For example, he may double the area of a figure, with a theoretical error under 1%, by translation of points on its perimeter through any combination of vertical and horizontal distances that sum algebraically to 3 db. Equal displacement in decibels along both axes preserves the shape of the figure. Translations that sum to 0 db., such as 3 db. upward and 3 db. to the left, result in a change of shape but not of area. The technique described is most successful with straight-sided figures, whose sides remain straight under these transformations.

---

[3] E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951, 560–561.

Conclusion. It is obvious that inaccuracy in construction or plotting, the need to estimate, round, and interpolate values in decibels, and the shortcomings of graphical methods in general, set limits to the usefulness of the antilog grid. Nevertheless, the grid is a convenient tool for many applications including those outlined and others that will undoubtedly suggest themselves to the user.

Washington University G. M. SHICKMAN

LINGUISTIC AND NON-LINGUISTIC PROCESSES IN SERIAL LEARNING AND RETROACTIVE INTERFERENCE

The results of a study on retroactive interference recently reported suggested that learning a list of conventional words by the method of serial anticipation may involve at least two distinct processes: (1) a *linguistic* process, by which subsections of the list are formed into highly stable, functionally integrated linguistic units; and (2) a *non-linguistic* process, by which (nonlinguistic) relationships between these linguistic units are formed.[1] In addition, the evidence seemed to indicate that while linguistic relationships form very rapidly, during the first few trials of anticipatory learning, the non-linguistic relationships develop only gradually during the remaining trials of anticipatory learning.[2] We noted, at that time, that one implication of these assumptions is that the curve of the serial position-error (*SPE*) for a *mixed* list of linguistic *and* non-linguistic relationships should be 'scalloped' in form, such that "(a) errors are maximal in the vicinity of those items which are at the boundary between linguistic fragments; and (b) errors should be minimal between items within each linguistic fragment."[3]

The present Note reports the results of an analysis of the *SPE* curves for those interpolated learning lists of this experiment which were, in fact, composed of mixed linguistic and non-linguistic relationships: viz. the *low* and the *intermediate* unity interpolated lists.[4]

---

* This work was supported, in part, by NSF Grant GB 1759, with the author as principal investigator.

[1] Sandor B. Brent, Linguistic and non-linguistic processes in learning and memory interference, this JOURNAL, 79, 1966, 181–194.

[2] For a more detailed description of methodology, see Brent, *op. cit.*, 182.

[3] Brent, *op. cit.*, 193, Item 2.

[4] For a more complete discussion of methods and materials, see Brent, *op. cit.*, 182. In the present analysis, only those *Ss* who received eight practice trials on the interpolated task were included, since this is the group for which the opportunity for inter-item variability in error rate was greatest.

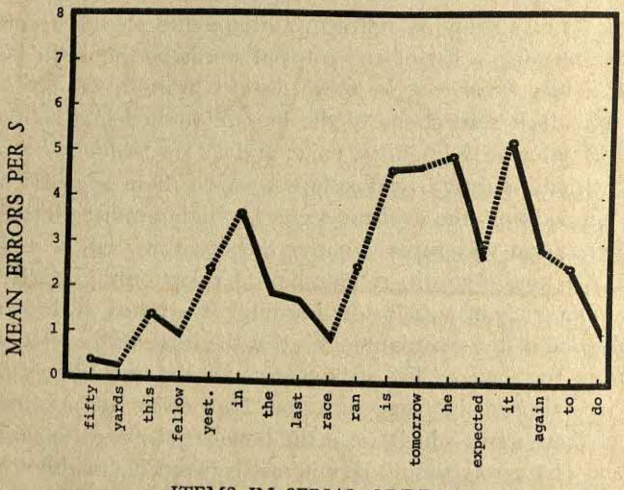
The results of the analysis for the *high* unity list are not reported here in detail

The *intermediate* unity list was hypothesized to contain 9 linguistic units: *fifty yards / this fellow / yesterday / in the last race / ran / is tomorrow / he expected / it again / to do.*[5]

The *low* unity list, on the other hand, was hypothesized to contain 11 linguistic units: *tomorrow / last yesterday / yards / it again / be / fellow / to do / is expected / race fifty / in this / the ran.*

According to hypothesis, the pattern of 'scalloping' of the *SPE* curves for each list should, in general correspond to the division between linguistic units based upon an earlier, purely linguistic, analysis.

Results. Figs. 1 and 2 show the *SPE* curves for the *intermediate* and *low* unity lists, respectively. In these figures, the *opened circle* denotes the first item in each hypothesized linguistic unit; the *solid line* corrects items within each unit; the *broken line* marks the transition from one unit to the next. (An isolated word is considered the first item in a linguistic unit containing only one item.)



ITEMS IN SERIAL ORDER

FIG. 1. *SPE* CURVE FOR INTERMEDIATE UNITY-LIST

---

since: (1) the *SPE* curve for this list did *not* show any noticeable 'scalloping'; and (2) the highest *mean* error rate per item over the entire list was less than 1.0. Both of these findings suggest the hypotheses that the two-sentence paragraph, which constituted the *high* unity list, was learned virtually completely after the first trial of anticipatory learning, and functioned as a single item of experience thereafter. (Cf Brent, *op. cit.,* 187.)

[5] Slash-marks have been introduced to indicate hypothetical linguistic subunits, based upon a linguistic analysis of the intermediate and low unity list (see, below, and *op. cit.,* 187). In actual stimulus-lists, these slash-marks did *not* appear.

Examination of these curves shows: (1) the transition *between* linguistic units, in general, is accompanied by an increment in error rate (indicated by the positive slopes of the broken lines); (2) the transition from one item to the next *within* each multi-item linguistic unit is, in general, accompanied by a decrement in error rate; and (3) the first item of each multi-item unit is, in general, the most difficult item within the unit to learn (opened-circles). Thus, on the basis of visual evidence, at least, the *SPE* curves for these *mixed* lists do appear to be "scalloped" in the predicted manner.

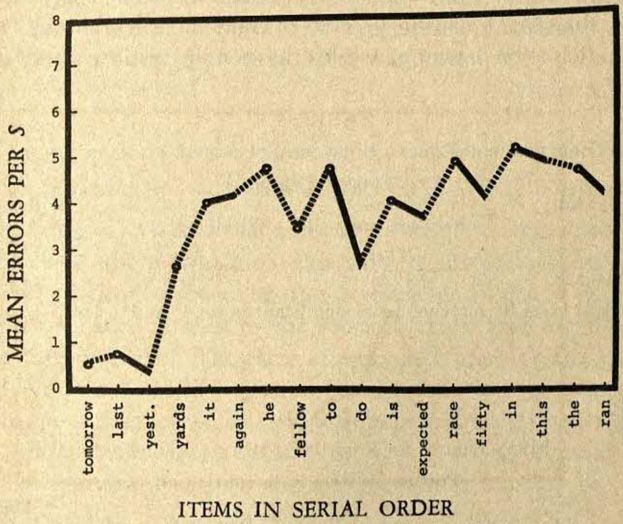


FIG. 2. *SPE* CURVE FOR LOW UNITY-LIST

To evaluate the reliability of these results, two types of statistical tests were performed: *(a) Between* vs. *within unit-slopes.* Different scores between successive items in each list were divided into two classes: those representing the transition from one item to the next *within* a linguistic unit (solid lines) and those representing the *transition* from one linguistic unit to the next (broken lines). Thus, each *S* received two error-scores: a mean *between-unit* difference-score and a mean *within-unit* difference-score. According to the hypothesis, the mean score between-units should be significantly larger than the mean score within-units. The results of this analysis, summarized in Table I, show that the differences were significant and in the predicted direction—*i.e.* that a transition across a boundary between linguistic units was, in general, associated with an increment in errors, while a transition between items within a unit was, in general, associated with a decrement in errors.

(b) First vs. *non-first items in a unit.* Items within each list were again divided

into two classes: (i) those which were the *first* item in a linguistic unit (opened circles on the graphs); and (ii) those which were *not-first* in a unit. Each *S* was again given two scores, representing his *mean* errors on each type of item. Results of this analysis, summarized in Table II, support the hypothesis that the first item within a linguistic unit tends, in general, to be more difficult to anticipate correctly than are the other items within a linguistic group, hence, again, that learning to anticipate the order of the groups themselves is more difficult than learning to anticipate the order of items within the groups.

The results of the present analysis provide additional support for the hypothesis that the linguistic process operates rapidly during the early trials of anticipation learning, while, the non-linguistic process operates gradually during later trials. They also provide support for the hypothesized division of these lists into linguistic units, based on a purely linguistic analysis. This support must, at this time, however, be taken as indicative rather than conclusive. Caution must be exercised for a number of reasons: (1) while the present analysis is independent of that which suggested the current hypothesis, the experimental data used in both the present and former analyses are identical; (2) the *SPE* curves reported here were derived from the interpolated learning phase of a retroactive interference task, hence the proactive effects of the original learning task *may* have been of decisive influence here (although, at present, we cannot see why this should have been the case); and, finally (3) since these lists

TABLE I

MEAN DIFFERENCE IN ERRORS BETWEEN SUCCESSIVE ITEMS IN EACH LIST

Level of Unity	Type of Difference		Difference
	Between-units	Within-units	
Low	+ .76	− .63	1.39**
Intermediate	+1.22	−1.06	2.28**

** One-tailed *t*-test for matched-pairs significant beyond the 1% level of confidence, *df*=13.

TABLE II

MEAN ERRORS ON EACH ITEM WITHIN EACH LIST

Level of Unity	Position of item in linguistic unit		Difference
	First	Not-first	
Low	3.88	3.30	.58**
Intermediate	3.36	1.89	1.47**

** One-tailed *t*-test for matched-pairs significant beyond the 1% level of confidence, *df*=13.

were not originally constructed to test these hypotheses, no systematic explanation of the effects of level of linguistic unity itself could be explored.

On the basis of the promising nature of the present analysis, an independent study of the effects of linguistic unity upon the structure of the serial position curve has been undertaken and is currently near completion. Analysis of partially completed results of this second study indicate that substantially the same "scalloping" effect occurs whenever the boundary between one linguistic unit and another is crossed, depending upon the level of linguistic unity being analyzed.

Wayne State University SANDOR R. BRENT

THE GAME OF JAM—AN ISOMORPH OF TIC-TAC-TOE

In psychological research on problem-solving, sometimes the game of Tic-Tac-Toe is employed, which, though very simple to learn and play, still offers sufficient problems to the investigator in that it is not at all clear what heuristics are used by the subjects, except avoiding the winning move of the opponent.[1] The same is apparently true for the isomorphic game of Number Scrabble, which is based on the fact that there exists a 3 × 3 magic square, of which rows, columns, and main diagonals add up to 15. This game is usually played by having players alternately select one of the remaining numbers from a pile which contains the numbers 1 to 9 at the outset. The player wins who first collects a triad of numbers adding up to 15.

Although the games are mathematically equivalent, they are likely to differ psychologically, because they require different sorts of information. Number Scrabble is a numerical game requiring addition and subtraction of numbers, whereas Tic-Tac-Toe requires a spatial representation, which can only be disposed of with some difficulty, even by fairly experienced players. On the other hand, it is very easy, though maybe not likely to happen, to have one's 'moves' in Number Scrabble determined by superimposing it on the Tic-Tac-Toe configuration, thus using the spatial magic square.

A comparison of strategies used in both games may well show marked differences. Such comparison, however, could possibly suffer from the easy superimposability of the two games, which in some cases might even amount to 'psychological isomorphism.' The same kind of problem will arise when we set out to study the effect of training in one game, upon playing the other, the so-called transfer of training.

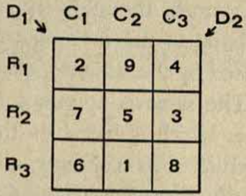
---

[1] This Note was prepared while the author was at Carnegie Institute of Technology, Pittsburgh, Pennsylvania, as recipient of a NATO-Science Fellowship, granted by the Netherlands Organization for the Advancement of Pure Research, Z. W. O.

The game to be described is also isomorphic with Tic-Tac-Toe; it is not superimposable spatially on the Tic-Tac-Toe board, and is less easily translated into Number Scrabble. These properties make it a welcome aid to the investigator who wants to study problems of subjective strategies in simple games.

The Game of JAM

The construction of the game of JAM is in fact a simple problem of projective geometry: Construct the dual of the game of Tic-Tac-Toe, and see whether this results in a sufficiently simple geometric configuration to make it a playable game. That is, no lines should intersect at points which are not proper vertices of the configuration (planar graph), while the configuration as a whole should have some symmetrical appeal.



	C1	C2	C3
R1	2	9	4
R2	7	5	3
R3	6	1	8

FIG. 1. THE MAGIC SQUARE REPRESENTATION OF NUMBER SCRABBLE

Construction turned out to be fairly easy and straightforward. The Tic-Tac-Toe game embodies nine vertices (cells) and eight lines (winning triads) related in the way shown in Fig. 1 and Table I. The elements of the dual problem are obtained by so interchanging the properties of lines and points that we now will have, for instance, one line (5) with four points on it (D_1, D_2, C_2, and R_2) instead of one point with four lines going through it. If we start with Line 5 and take this as the axis of symmetry, we easily obtain the configuration shown in Fig. 2. It is clear that this arrangement is only symmetric along its vertical axis, a consequence of the asymmetry of the numbers of points and lines in the two problems. Apart from rubber sheet transformations and a few trivial shifts up and down along Line 5, of Points C_2 and R_2, this is the most simple configuration satisfying all our requirements.

The Rules of JAM

The game has been called JAM because, among other reasons, it may be visualized as a game in which the players block roads (lines) and in doing so try to jam crossings (vertices). Each of the two opponents alternately selects one road, until one of them has occupied all three roads leading into a crossing. Legal moves are only those which occupy the whole of a road, irrespective of the number of crossings it contains. So in Fig. 2, D_1C_2, R_2C_3, or C_1R_3 are legal, whereas C_1D_2, or D_2R_2 are not. An alternative game, which is more positively oriented to traffic safety, is that players try to avoid jamming and that he who is first forced into it is fined.

TABLE I

PLAN OF THE VARIOUS GAMES

Tic-Tac-Toe	Through point	Go lines
JAM	On line	Lie Points
	5	D_1, D_2, C_2, R_2
	2	D_1, C_1, R_1
	4	D_2, C_3, R_1
	6	D_2, C_1, R_3
	8	D_1, C_3, R_3
	1	C_2, R_3
	3	C_3, R_2
	7	C_1, R_2
	9	C_2, R_1

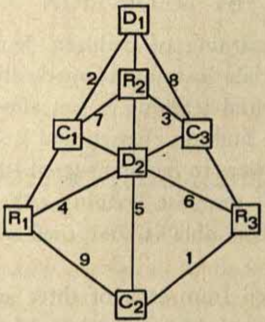


FIG. 2. OUTLINE OF THE BOARD USED IN PLAYING JAM

The game of JAM is most easily played if a board is prepared in some light color, such as yellow or green and a thin sheet of celluloid or perspex is placed over it. Players then can mark the roads they choose with blue and red marking pens. Their markings can easily be wiped off after each game.

The equivalence of Tic-Tac-Toe, Number Scrabble, and JAM, will be immediately clear. Superimposing Number Scrabble and JAM might be done by giving the lines numerical values (as in Fig. 2) but this procedure is certainly less obvious than in the case of Tic-Tac-Toe. Furthermore, spatial superimposition of Tic-Tac-Toe and JAM is by definition impossible. Without writing the relations down explicitly, even the mapping of the two games onto one another is not easily accomplished.

Although thus far no systematic comparison between the three games has been undertaken, it appears off hand that JAM is the easiest to learn. This impression may have been colored, however, by the circumstance that there will be hardly a subject in our western culture who has not been exposed to Tic-Tac-Toe.

Though the author has not been able to find a reference in literature pointing to an earlier construction of the game of JAM presented here, it would be very surprising if it had never before been played. Information pointing to the contrary will be appreciated.

Institute for Perception RVO-TNO JOHN A. MICHON
Soesterberg, Netherlands

SERIAL LEARNING OF CHILDREN AS FACILITATED BY MEDIATION

In a study on verbal learning of children, Jensen and Rohwer found that whereas paired-associate learning is markedly facilitated by instructions to use mediators, serial learning is not affected by this procedure.[1] In spite of their negative findings, Jensen and Rohwer had proposed that if mediating instructions were to facilitate serial learning, this effect would be due to a facilitation of response learning rather than to facilitation of serial learning *per se*.[2] This note is concerned with a test of the Jensen and Rohwer hypothesis.

Twenty *S*s were selected from each of three age-levels; 7, 12, and 17 yr. of age. The 7- and 12-yr. olds were selected from an elementary school which was in the same district as the secondary school from which the 17-yr. olds were selected. The materials consisted of 10 colored pictures of common objects taken from a preprimer workbook used in Vancouver schools. The pictures were of the following objects: apple, ball, tree, sun, flower, fish, hat, table, boat, and chair. Each picture was mounted on a 3 × 3-in. cardboard.

The 20 *S*s at each age-level were randomly assigned to two groups;

---

[1] A. R. Jensen and W. D. Rohwer, Jr., Syntactical mediation of serial and paired associate learning as a function of age, *Child Devel.,* 36, 1965, 601–608.

[2] Jensen and Rohwer, *op. cit.,* 606.

one group (Group *N*) had to name the pictures, while the other group (Group *M*) had to make up sentences containing the names of each successive pair of pictures, thereby chaining words 1 and 2, 2 and 3, etc. Therefore, 10 *S*s at each age-level were tested under each of the conditions of verbalization. The same procedure was used at all age-levels. Each *S* received pretraining to a criterion of one perfect recall, and was then tested for serial learning during the same session. The pretraining procedure consisted of showing the pictures to *S* and asking him to name each picture. A different serial order was used on each pretraining trial.

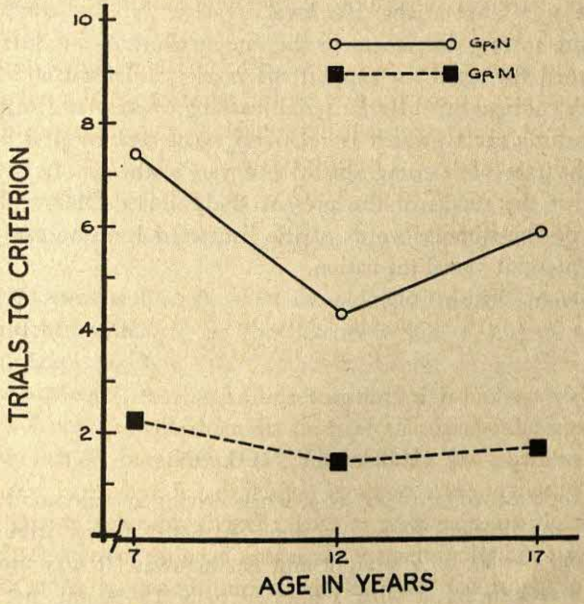


FIG. 1. MEAN NUMBER OF TRIALS TO LEARN A 10 ITEM LIST

During the test for serial learning, the *S*s in Group *N* named each picture as it was presented on the study-trial, and anticipated each picture on subsequent trials. A serial order different from the ones used for pretraining was used on the test-trials. After *S* made a response on the test-trial, *E* showed the picture which should have been named in that serial position. This was repeated until *S* was able to recall the entire list in correct serial position, or until *S*s had completed 15 trials.

The *S*s in Group *M* were required to chain the names of the successive pair of pictures by constructing a sentence to this effect. They were first given a study-trial in which the pictures appeared in a different serial order than they appeared during pretraining. The study- and test-trials consisted of exposing *S*s to pairs of pictures. The procedure was repeated to the same criterion required of *S*s in Group *N*.

The effects of instructions of verbalization and age of *S* on the trials to criterion measure is shown in Fig. 1. The performance of the *S*s in Group *M* was superior to those in Group *N* ($F = 9.80$, $p < 0.01$). The comparisons among the age-groups did not, however, yield differences which were significant at the 5% level.

Our failure to find differences in the rate at which *S*s of different age-groups learned the serial list support the results of Jenson and Rowher's study.[3] They contend that because serial learning constitutes a fundamental kind of learning ability which is relatively unaffected by past verbal experience, the rate of learning should not vary with age. It is apparent, however, that the results of the present study indicate that when *S*s are pretrained on the stimulus-words of the list, serial learning can be facilitated by sentential verbal mediation.

University of British Columbia

RODERICK WONG
DIANA MITCHELL
GAIL DESLAURIERS

ILLUSORY MOTION

During the past winter (1965-66), while driving an automobile at night, I observed the illusions of motion reported here.[1] I was driving east at about 45 mi. per hr. on a straight and flat highway. It was snowing and the wind was blowing at what I guessed to be 15–30 mi. per hr. from my back. Under the illumination of the headlights, I could easily see the snow blowing on the highway in front of and along the side of the moving car.

After driving for a short time, I became slightly nauseous—car-sick, which I rarely become. I suddenly became aware that I was seeing something similar to the illusion described by Du Mas, who named it the "radial illusion."[2] The falling snowflakes, which were in fact blowing in the same direction as the automobile was travelling, appeared to be moving toward the windshield as if the wind were blowing them in the opposite direction, that is from east to west, and yet I could also see snow blowing from west to east on the surface of the highway. I found my attention alternating between these two incongruent perceptions. These conflicting perceptions, I believe, were, at least partially, associated with my nausea.

---

[3] Jensen and Rohwer, *op. cit.,* 606.

[1] For another illusion, also observed while in a moving automobile, see K. M. Dallenbach, The elastic effect: An optical illusion of expansion, this JOURNAL, 66, 1953, 634–636; Recurrent images, *ibid.,* 35, 1925, 155.

[2] F. M. Du Mas, The radial illusion, this JOURNAL, 69, 1956, 118–121.

After about 10 min. of being aware of this perceptual incongruity, I then observed a most unusual sequel. The directionality (actual) of the blowing snow on the highway gradually became less distinct, until it seemed that the snow, on the highway was blowing from east to west. This perceived directional reversal of actual moving stimuli undoubtedly was the result of many contributing factors which research alone may reveal, but I believe that the urgency of the driving situation forced attention away from the lower periphery of the visual field where part of the incongruity could be observed. It may be as Graham has stated, "the threshold for peripheral movement is higher than the threshold for central vision".[3]

Motion parallax, as discussed by Gibson, Olum, and Rosenblatt, may also have an important bearing on the described illusion and its unexpected (by me) perceptual sequel.[4] The apparent reversal of the movement may bear some similarity to Korte's observations [5] He found, under certain conditions of intensity of illumination, size of objects, distance between objects, and length of pause, that reports of reversed movement can be obtained.

In my opinion, the subjective discomfort and the unexpected perceptual reversal of actual motion were the result of perceiving the simultaneous blowing of snow in opposite directions. Had the wind been blowing at a markedly different angle, it is doubtful that these perceptions of movement would have been perceived as being incongruent. Further study of these phenomena seems warranted.

University of North Dakota LEROY A. STONE

---

[3] C. H. Graham, Visual perception, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 895.

[4] J. J. Gibson, Paul Olum, and Frank Rosenblatt, Parallax and perspective during aircraft landings, this JOURNAL, 68, 1955, 372–385.

[5] Adolf Korte, Kinematoskopische Untersuchungen, *Z. Psychol.*, 72, 1915, 193–296.

SEVENTY-FOURTH ANNUAL CONVENTION OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

The Seventh-Fourth Annual Convention of the American Psychological Association was held this past year on September 2 to 6 in New York City. As has become customary, this meeting taxed the facilities of the four large hotels in which it was held, and taxed equally the stamina of any psychologist who attempted to take part in more than a very small portion of the total underaking.

In the past two years, the length of the meeting has been reduced deliberately from seven to five days simply because members did not attend for the full week under the old plan. In consequence, far more activities are scheduled at one time. The count for the busiest hours runs over twenty simultaneously scheduled events. Special features, such as the address of President Nicholas Hobbs, have been relegated to evening hours.

One other innovation of the past two years should be noted. This is the publication of relatively long summaries of papers before the Convention begins. This was tried in 1965 for the first time with support from the National Science Foundation, and was extended somewhat in 1966. The publication goes under the title of Proceedings of the Annual Convention. There is hope, and some favorable evidence, that this prompt and limited publication may relieve pressure on the more conventional journals. There is also the expectation that there may be a corresponding change in the pattern of paper-reading toward questions and discussions instead of reciting experimental data.

The annual awards and the speeches by the recipients provide a brief spotlight for seeing psychology whole. The Distinguished Scientific Contribution Awards were made to Richard L. Solomon, Clarence H. Graham, and Nancy Bayley. The Gold Medal of the American Psychological Foundation was presented to Karl M. Dallenbach. A new award, the Richardson Creativity Award, was given to Joy Paul Guilford. The writer of this Note is sure that readers will join him in felicitating Professor Dallenbach. He has been editor and publisher of the JOURNAL for over 41 years. The Gold Medal affords him well-merited recognition for his many and varied services to our profession.

The affairs of the APA were handled earnestly and yet expeditiously in two meetings of the Council of Representatives. The persistent issue facing the members of the Council is how to be all things to all men, and this with limited resources. Publications remain a little over half of the dollar activities of the Association, while the many and varied activities that touch on public relations and public policy consume the larger measure of staff time. The focus of this year's discussion was a plan for overhauling the structure of the Council itself. The debate on that issue continues. The Treasurer reported that the new $3,000,000 Headquarters building was fully occupied (including two floors used by the APA).

It is doubtful that any great scientific discoveries were reported. None came to your reporter's ears. What is impressive, rather, is the involvement by psychologists in social issues and practical affairs. There was a good deal of earnest talk about Vietnam, about new methods in psychotherapy that bring mental health from the hospital into the community, about mental retardation as a problem for research and correction. Today, hundreds of people are working in areas which a few years ago were the provinces of a lone pioneer. We can take pride in what is being done. With intelligence and hard work and good-will, tomorrow will be better.

Harvard University EDWIN B. NEWMAN

ERRATUM

In the preparation of the December, 1966, number (Volume 79, p. 517) Line 14 from that page was dropped and replaced by Line 9. The last three lines on this page should read:

more stable in visual problems, which usually yield less rapid improvements both within and between reversals. The work was done (again out of interest in generality) with pigeons and with rats—the pigeons

ED.

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Mood and Personality. By ALDEN E. WESSMAN and DAVID F. RICKS. New York, Holt, Rinehart and Winston, 1966. Pp. XIV, 317. $5.95.

The concept of mood and the phenomena, both covert and overt, to which mood terms are applied have received increasing interest in the past decade. This clearly written and well-organized book is an important contribution to this trend. It contains informative reviews of recent studies of normal mood, of pathological mood, and of human variability. It fully describes the authors' own two studies of mood and personality in college students, from which emerged a vast amount of data which they have dominated with some success. It also makes available several interesting new paper-and-pencil tests and the theoretical and empirical contexts within which they may be both used and hopefully improved by others; in particular, there are a set of 16 Personal Feeling Scales and a set of 60 items categorized with respect to successful and unsuccessful outcome of each of the first six of Erik Erikson's eight psychosocial stages.

Their major study was carried out in conjunction with a research project directed by Henry Murray, a project in which Couch, Keniston, McNamara, and Nielson were also participating. From a group of 70 Harvard sophomore volunteers, 27 men were selected for a three-year "intensive assessment study." In the fall of their senior year, 18 of these students were invited by Wessman and Ricks to fill out the set of Personal Feeling Scales just before retiring on each of 42 successive nights. Complete records were obtained from 13; nearly complete, from 4 others. Each was also asked to self-administer three other tests twice during the last two weeks of the school period, once when highly elated and once when very depressed: A Q-sort of the Eriksonian items for both self- and ideal-description, the Rosenzweig *P-F* Test and the Cattell *16 PF* Test. In the minor study, done earlier in collaboration with Mary McIlvaine Tyl, 21 Radcliffe students went through a similar six-week procedure; there were only 11 Personal Feeling Scales, the *16 PF* Test was omitted, and the Q-sort utilized not Eriksonian items but 45 relating to introception, social interaction and work. No other information relevant to personality was available for these women, but the larger project provided for the men's life histories, diaries, interview data, and scores on at least 365 personality variables.

In the five chapters which constitute the main contribution of the book, the authors apply these data to the consideration of five general questions:

(1) How did these subjects differ when elated and depressed? Here, the test-results obtained when the students felt they were very depressed are compared with the results obtained in high elation. When depressed, both men and women showed loss of self-esteem, greater discrepancy between self- and ideal-descriptions, greater feelings of social withdrawal and of self-preoccupation and, on the *PF* Test, contrary to expectation, an increase in extrapunitive rather than in intropunitive responses. Scores on the *16 PF* Test were fairly stable except for a significant decrease, with depression, in Factor *F*, Surgency.

(2) How happy were these college students and how did their levels of hap-

piness vary from day to day? From the beginning, the item of greatest interest to the authors in their set of Personal Feelings Scales was a 10-point bipolar scale of elation vs. depression, which they also refer to as hedonic level or as "relative happiness-unhappiness." As with the 15 other personal feeling scales, the subject recorded nightly on this scale his estimates of his highest point (peak) during the day, his trough, and his average level or "over-all summary" of the day. As expected, they tended to average just slightly on the happy side of the neutral point of the scale. Contrary to some earlier studies, both day-to-day and within-day variability were independent of mean level of elation vs. depression, an important finding which led to the decision to make separate analyses of the personality correlates of the happy and the unhappy and of the variable and the stable. The results of four factor analyses were also interpreted as consonant with this decision; for both men and women, a centroid factor accounting for more than 50% of the common variance, extracted from the matrix of correlations of the means of daily self-ratings, showed high loadings for all personal feeling scales; similarly for both groups, a first large factor, accounting for over 80% of the variance, extracted from the matrix of correlations computed from the standard deviations of the daily ratings, showed high loadings for all scales.

(3) Did these individuals differ in the way in which other aspects of mood were associated with changes in elation and depression? With an implied admission that the previous four analyses did not provide adequate examination of the factorial structure of their set of Personal Feelings Scales and without further attention to ways to improve that examination, the authors decided to factor analyze 41 different correlation matrices, one for each of the 17 men and 21 women who had completed from 30 to 42 daily records, as well as for three other women with almost 30 records. Since elation-depression was for them the most salient variable, the first centroid factor in each matrix "was always rotated through average Elation vs. Depression, and the other factors were then rotated to an orthogonal approximation of simple structure" (p. 73). The procedure was designed to discover in each subject (1) which other scales were highly loaded on a first factor which had average elation-depression as the marker variable; and (2) how many factors there were in each rotated matrix which individually accounted for 10 percentage or more of the common variance. How does one so dominate the results yielded by 41 factored matrices that meaningful statements can be formulated? The authors' method is first to survey the results for evidence as to which scales, if any, tend to covary with elation-depression in all subjects, or tend to be independent, or tend to fall on both the first factor and on others: They then sort and examine the matrices for men vs. women, the happy vs. the unhappy, and the stable vs. the variable. Finally, they discuss the matrix of each of five men in the light of the extensive knowledge available about them from the three-year study: these five were their most happy man, most unhappy, most stable, most variable, and one man who kept a particularly full diary throughout his college years. The whole effort is of interest because Wessman and Ricks are attempting one possible way of combining the ideographic with the nomothetic approach to the study of personality. They frankly discuss (or demonstrate) some of the advantages, difficulties and interpretative leaps and pratfalls such an approach may afford. When we psychologists choose to communicate with computers as well as with subjects and clinicians, most of us, like the present authors, have to depend on helpful colleagues as mediators between us and the machine. The results, especially early in the game, are rarely satisfactory.

(4) What are the personality characteristics associated with happiness and unhappiness? When the mean daily average of elation-depression in the 17 men was correlated with the 365 personality variables, 68 coefficients reached a significant value (0.48): 0.71 with happiness level estimated six months earlier by clinicians working in the assessment study; —0.83 with the Depression Scale, and —0.64 with the *F* scale of the *MMPI*; 0.48 with valuation of happiness as a life goal, 0.43 with four-year college grade record, etc. It appears that average level of happiness is an important characteristic of the college student. It is also important to note that self-estimates of it are relatively free of the influence of response set and that there is new evidence, indirect in the present work and more direct in later work, that *only one* sincere estimate of typical level of elation-depression may serve almost as well as an index based on 42 daily estimates.

(5) What are the personality characteristics associated with stability and with variability of personal feelings? In studying correlates of affective variability, the authors' task was speeded by the previously mentioned findings that variability of scores on the scales was independent of mean score and that the standard deviations of mean scores on all scales were highly correlated. A further finding also simplified their task: within-day variability on the E-D scale (as measured by mean daily range between trough and peak) was highly correlated with the standard deviations of mean peak, trough, and average E-D scores. When all four variability measures were correlated with the 365 personality variables, at least 61 variables were related to one or more of the variability measures; more than half (31) were with mean daily range. By contrast with 68 significant correlations found with one measure of mean level of elation-depression, it appears that measures of affective variability are less generally related to the kinds of personality measures available for these men. The 61 variables are grouped into three rough categories: (1) variable men give many more indications of emotional openness; (2) stable men show greater maturity, self-sufficiency and sense of identity; (3) in temperament, variable men are more energetic, fast-paced and cyclothymic.

This summary includes only a small sample of the results of these studies and of the hypotheses and interpretations which are presented. Since the book and the studies it reports stand at the intersection of many current sets of research activities aimed at increasing our understanding of personality, particularly the personality of the college student, some attention or at least mention should be given to their flaws, whether corrigible or incorrigible.

Flaws that are incorrigible, at least in the near future, appear in this and in all current attempts to find a language for dealing with the muddled domain of affect, feeling, emotion, mood, and temperament. Since capricious substitution of undefined terms for each other occurs here no more frequently than it does throughout the relevant literature, we can forgive the authors for reflecting the muddle. But we must also establish some discriminations that do affect the interpretation of their main results. In the first place, I think their procedures have yielded more information about temperament than about mood, although this is not an all vs. nothing distribution. Temperament is disposition toward this or that mood. To obtain a mean of 42 records of estimated average daily level of that dimension of mood called elation-depression tells us more about one aspect of temperament than about elation and depression, about trait rather than state. To obtain the standard deviation of those 42 averages again tells us more about an aspect of temperament ("moodiness") than about elation and depression. This discrimination is not just a semantic nicety but calls attention to the importance of relating size of temporal unit to the choice of research procedures and of appropriate hypotheses. Analysis of a brick calls for strategies different from those appropriate for analysis of a brick wall. If the authors had really been interested in identifying the factorial structure of affective states in mood, they would have factor analyzed a set of data based on their minimal temporal unit—one day. Analysis of each set for each of several days would have been of still more interest—an actual day like Thanksgiving or the day of the Harvard-Yale football game—or a synthesized day, with records from each subject on *his* big exam day or "high" or "low" day. Such analyses would be subject to some of the same methodological criticisms which can be made of the 42-day average analysis, where the N in a 51 × 51 matrix was only 17. But surely an examination of the characteristics of the Personal Feeling Scales could easily and should have included at least one administration to hundreds of subjects—and almost certainly would have shown greater independence among these scales when used to measure mood rather than temperament.

Another token of the muddle in this domain was the initial assumption that the Personal Feeling Scales would reflect dimensions of feeling more adequately if set up as bipolar rather than unipolar scales. They made this assumption before the recent appearance of a considerable body of evidence, based on factor analytic studies of reports of average mood level "for the past week" as well as reports of momentary mood, that mood scales should be unipolar, ranging from a neutral point to one extreme point. Fluctuations in any "positive" dimension of mood may be quite independent of fluctuations in any or all "negative" dimensions. Bipolar scales obscure this independence. Some of the Wessman and Ricks data suggest that their subjects did perceive the positive and negative halves of their scales as somewhat independent as they were checking their daily peaks and troughs. As example, the correlation between mean daily peak and mean daily trough was approximately zero and the two sets of means develop different patterns of correlations with the other variables.

Another assumption stemming from inadequacy of terminology in this domain influenced the authors' final interpretations as well as the nature of their results. In their original Personal Feeling Scales, as given to the subjects, the elation-depression scale, and only this scale, was given the label "Mood," reflecting a traditional view of mood which is most common in the psychiatric literature. As they reviewed more of the recent studies of mood, they revised their own view of mood to encompass dimensions other than elation-depression. But all they had available as possible indices of some of these other dimensions were the 15 other Personal Feeling Scales, which were not originally formulated on the basis of hypotheses about mood dimensions as such. I do not know of any study which combines the Wessman-Ricks scales with scales for the mood dimensions identified in other recent studies. A number of their scales (harmony vs. anger, sociability vs. withdrawal, tranquility vs. anxiety), particularly if set up as unipolar scales, would probably correlate with marker variables from other studies. Others would not; are probably not unitary and possibly are poorly defined. Moreover, the fact that only one scale is used to sample each dimension makes it very difficult to interpret their many factored matrices and to accept their conclusion that "One's general hedonic level or over-all happiness-unhappiness was broadly indicative of all one's affective experience (p.

245)." Finally, even though in the individual matrices the first centroid "factor was always rotated through average Elation vs. Depression (p. 73)," we find that, because of difficulties not described in the book, the loading of this marker variable on the first rotated factor fell to values as low as 0.52 (p. 295).

An example of the difficulties attendant upon the use of a factored matrix in the interpretation of individual personality is provided by the authors' index of affective complexity (p. 73), which is simply the number of factors each of which accounts for 10% or more of the common variance in the individual's matrix. Since this index correlates significantly (0.57) in the men with their stability, as measured by the standard deviation of the daily average report on the elation-depression scale, the authors conclude that stable men "had much more differentiated and complex affective variation than the variable men" (p. 24), possibly arising from "a detached, analytical, isolation approach to moods, with clear awareness and fine differentiation of the characteristics of different feelings" (p. 73). This conclusion is in conflict with their later assertion that the stable person is a "relatively repressed person (who) is inclined to avoid reporting even those feelings that force their way into his conscious life" (p. 186). Moreover, the conclusion should have been suspended until it was determined whether this index (the number of factors in a matrix) was largely due to simple restriction of range in the stable, leading both to lower intercorrelations and to the subsequent absence of a large first factor.

Despite these flaws, I found the book interesting and exciting. It should be read by everyone interested in happiness, in mood, in variability, in the personality of college students, or in the integration of clinical and statistical approaches to personality.

University of Rochester VINCENT NOWLIS

The Senses Considered as Perceptual Systems. By JAMES J. GIBSON. Boston, Houghton Mifflin Co., 1966. Pp. xv, 335. $7.50.

How James Angell and the Chicago group should have welcomed this book sixty years ago, the first truly functional psychology of sense-perception that has ever been written. The satisfaction of that need, which would have coincided with Gibson's birth, might even have defended Chicago's Dewey-generated faith from the lethal thrust of behaviorism. Now the resolution comes long after it was first wanted, novel, difficult because its conceptualization is so strange, certainly not yet quite fool-proof, yet a great improvement on "the ancients," if we may so style all pre-Gibsonites as Descartes did all pre-Cartesians in order to confound them. Certainly Gibson misses few opportunities to belabor his ancients as he unfolds his new evangel. If he seems to have forgotten that the history of thought is continuous and that the firm foundation of every new present lies in its past, nevertheless we may ask why he should not be caught in the motivational predicament of most originators who are overdetermined by their enthusiasms.

A better title for the book would be "The Senses as Percepthal Systems," and a still better one simply "Perceptual Systems." A more analytical title would have been: "Perception: How Ecological Information is Obtained," thus bringing out at the very start the three key words that make the book what it is: *obtain, ecology, information;* but such a title would scare the reader off, or, if it starts him at all, leave him puzzled until he has read a quarter of the book.

Well, here is the core of this functional psychology of perception, of Gibson's perceptual ecology, as well as I can put a book's worth in a few words.

The organism (more later about who is really having these perceptions) lives out its life immersed in a sea of ambient energy (*ambient* means going all around: *ambi,* around; *ire,* to go—if you don't know, as I didn't) which is differentiated or structured. The organism is active. It wants to know what is going on outside because that is the way evolution made it. A quiescent organism waiting around to be impressed by stimulation would not be here, would not have survived in life's struggle. So the organism is active, its construction is dynamic, and it "hunts" for stimulation that will give it security or command over its environment. (Gibson speaks mostly of *obtaining* information, but I think *hunting* is the better word. Woodsmen *hunt,* but so also do electric generators in finding the right speed for the right frequency, and so do compasses hunt for the north, now too far east, then too far west; even history hunts for truth, back and forth, but nobody has said that until here right now. See Gibson, pp. 83, 271, 303, 320.) This continual adjustment of the organism to the ambient flux of energy that is its environment belongs in the province of perceptual ecology (*eikos,* a house, what surrounds you, is ambient to you), and the organism by the hunting finds constancies in its environment, finds among the parts invariances that remain constant under the transformations that the variable environment and its own perpetual searching constantly provide. Size-constancy is a good example of such an invariant, for perceived size is (actual angular size × actual distance), which, being invariant under the transformation of distance, establishes the perceived size as an object. The organism gains security and command as it detects these structural invariances (objects that stay constant) in its environment. Remember: in this case the object, an invariance, is a product of two environmental quantities.

You don't have five senses any more; you have four classes of activity, four kinds of hunting: listening, touching, tasting-smelling, and looking. The environment, differentiated into objects, can be said to be stocked with information, and information *is* differentiation of stimulation. The process of obtaining stimulation, which is what perception is, of hunting for invariances, is a lively one. The organism is not stamped by a stimulus with an image; it goes out after information, and its search constitutes an operating perpetual system organized in both time and simultaneity across what we ancients used to call the various senses. Take looking. This is no business of structured light's forming an image on the retina which so gets projected on the cortex, where a center, assuming the role of Descartes' soul and of Johannes Müller's hominoid sensorium, accepts the imposition and knows what it is and whence it came. It is true that ambient light is structured and the active perceptual system hunts out the useful invariances by scanning, moving the eyes, the head, the body, as the successive retinal excitations change to lay hold of information which is not formed by the system, for it is there already in the environment. In short, the system is out shopping for ready-mades, though it may have to make some alterations at home where subjective factors supervene upon the objective. If all this does not make sense, then you must study Gibson's book. Do not merely try to read it; that won't help you. Study it.

So perceptual ecology consists in understanding how a perceptual system can gather information in the form of invariances and bring it in. Then what does the perceptual system do with it? It can not give it to a soul or a hominoid sensorium or even to a sensory center endowed psychophysically but mysteriously with consciousness. Gibson would have none of that, yet all through Gibson there is a hint of a ghost, just off stage, wanting the information, keeping the perceptual system at work, hunting, hunting. Now it is not Gibson who talks about "the organism." It is I. The organism is a substantial vicar for Gibson's ghost, and it is to protect Gibson's ghostly huntsman from disgracing itself in science by seeming to have a free will what I said that it was evolution and not a soul that kept it going.

Is the ghost conscious? At first I had the impression that it is. "Animals and men do perceive and respond to the permanent properties of the environment" (p. 3). "The active observer gets invariant prceptions" (p. 3). "Animals and men can select and enhance the stimuli" (p. 32). And then much later: "An expert perceiver develops a highly economical strategy of perception" (p. 309). "The individual . . . has to detect the world from [the] invariant properties in the flux of energy" (p. 319). But then on the next to the last page of the book: "A perceptual system hunts for a state of what we call 'clarity' " (p. 320). These are the phrases we use when we are talking about consciousness and free choice, but I do not believe that Gibson's perceiving organism is conscious as such, or meant by Gibson to be free. I think he would say that this is a foolish question, and indeed he does escape from this vocabulary when he turns to Schlosberg's word *register* (p. 37 *et seq.*) for what happens to the informational input. "To register," if I understand the meaning, is to affect the organism without the organism's necessarily knowing that it is affected, and the advantage of the word is that it denies or ignores consciousness and thus also free choice.

So it comes about by way of information-theory that we have at the present time a new kind of "empty organism," one that exhibits causes without direct connection between cause and effect, like Skinner's empty Humian correlations between stimulus and response that are causal without touching. An engineer said the other day of monaural auditory localization: "If the sound comes from below, it is double, the direct sound to the ear-canal and an echo from the pinna a few microseconds later, and that's information to your ear, and so your computer knows where the sound comes from." Your computer! That computer is hominoid like Müller's sensorium. It is enough to note that the information, the differentiation, is received. Perhaps Gibson's ghost is a computer, if information is enough, if it does not have to be processed.

It may be that Gibson has some such view, that ghosts can not be localized nor do they obey the law of causality. He distinguishes information *about,* which is equivalent to perception *of,* from information as structure, which is, of all things, discrimination (p. 245). The latter is directly registered, which seems to me to mean that it gets to the hominoid computer and that is all you need to know about it.

As to causality, Gibson tells us that the perceptual composite, its parts distributed in both simultaneity and time, need not be cumulated into a simultaneous whole in order to function; it is nonsense, he says, to think that successive events have to be preserved as traces for the sequence to act as a totality (p. 251). Well, Newton got away with action at a distance, preposterous as it seemed, and perhaps Gibson could ultimately win our assent to a temporal vacuum between a cause and its effect, but it would require revolutionary thinking.

Of course, I think I know how to solve Gibson's problem. Information *is* differentiation. Get rid of the ambiguous word *information* and substitute *differentiation.* The environment is differentiated in what is not merely a spatial pattern. The invariances are in it as relationships, and they give you the objects. These structures are picked out of the environment by the receptor systems, which may modify them by screening out some differentiation and also by putting some new patterning in. Objective entropy *plus* subjective supplementation. This composite is thus processed by the nervous system, with loss here and gain there, and it runs on into behavior, or else one would never learn of its existence, for there has to be some sort of report, verbal or instrumental, for observation to occur. The processed differentiation actually comes out as discriminative behavior, and right there the scientist picks it up as information, for only then can he see the whole process as a fact or a datum. (Cf. *Psychol. Rev.,* 44, 1937, 445-461.) In that account you have straight causality without action at a distance, with no misleading dynamic words like the hunting of the perceptual systems; and activity, when it does come in, is accounted for by dynamic mechanisms that evolution, which is no ghost, has produced.

Criticism takes so much space and praise needs so little that all reviews are unbalanced. There are scores of good things in this book. For instance, there is an enormous amount of interesting information about the evolution of animals in achieving efficient perception. Does everyone know that predatory animals have eyes in the fronts of their heads, along with necks to move the head around, a set-up which enables them to fixate their prey and go to dinner, whereas the preyed-on animals have eyes on the sides of their heads, with little or no overlapping vision, but more nearly a panoramic vision, which can keep guard on all sides? It is common knowledge that a rabbit, with an eye on each side, has more than panoramic vision, with a little overlap of the two visual fields in front and also, if the eyes are moved back, at the rear? This book is chock full of interesting facts of that sort.

The figures are magnificent. How Gibson must have worked to get them and prepare them for this perfect publication!

But the index is terrible. Someone shirked there. There are only about 250 entries, whereas Irvin Rock's slightly smaller current book on perception has about 4500 entries. Of course there is such a thing as over-indexing (Index: "Lord Chief Justice, his great mind." Text referred to: "The Lord Chief Justice said he had a great mind to hang the man").

Semantics keeps making trouble. Of subception, Gibson says: "To say that one can perceive in order not to perceive is a logical contradiction" (p. 291). Not if you've got two meanings for perceive, like *register* and *be aware of.* If a man after skull fracture has a memory that lasts only two seconds, is he unconscious? He can react within the two seconds, but he can not introspect unless every reflex and reaction counts as introspection. You look down a long list for the name you want to find and it catches your eye the instant you come to it, so you had to be aware of all the preceeding names in order not to notice them, but you can recall or recognize none of them.

And what about Gibson's list of quotations on the last page of the front matter? They are from Aristotle, Hobbes, Johannes Müller, and Helmholtz, presenting what came to be called the specific energy of nerves, and then the last quotation is from the philosopher J. L. Austin, who said: "There is nothing so plain boring as the constant repetition of assertions that are not true." No reference is given, but I don't believe that Austin was talking about specific nerve energies. It is Gibson who is bored, who believes these ancients wrong, and he puts Austin in as an *argumentum ad verecundiam.* Actually the first four statements were true within the limitations of the Zeitgeist in which they were uttered. There is an enormous difference between Müller and the Greeks. The eidola from objects never get to the brain. The brain, if I may skip a century for my words, gets only coded information. That was as important a discovery as Glisson's finding of muscular irritability in 1677.

It is doubtful, I suppose, that Gibson, with his many rude iconoclasms, can please an ancient like me, and indeed, if revolutions did not swing too far over, they would not carry enthusiastic youth with them nor get the new ideas well enough assimilated for the ideas eventually to be modified, not reversed ordinarily, but redirected, by the next generation. Certainly Gibson's volume is the most original work we have had in the field of sense-perception for a long, long time. Some of it, I think—the functional amplification of perceptual system—may never be reversed at all, but the details invite dissent, and the progress of civilization depends, of course, on the interaction of dissents.

Harvard University

EDWIN G. BORING

Sleep, Dreams, and Arousal. By EDWARD J. MURRAY, New York, Appleton-Century-Crofts, 1965. Pp. 407, $7.50.

The Psychology of Sleep. By DAVID FOULKES, New York, Charles Scribner's Sons, 1966. Pp. 265, $6.96.

Enough books on sleep and waking have appeared in the sixties to cause authors and publishers real problems in selecting novel titles. Perhaps this accounts for the inaccurate labeling of these two works. Murray's book might more validly have been called *An Analysis of the Sleep Motive,* while Foulkes's work could be labeled simply *Rapid Eye-Movements and Dream Reports.* The latter book is brief, little more than monograph length, and is devoted almost exclusively to that difficultly private experience we call dreaming. Murray, on the other hand, has prepared a much broader and more intensive volume ranging over the entire psychological and psychophysiological gamut of sleeping and waking. He dwells on dreaming for less than fifty pages in all. The rest is addressed to such topics as sleep, learning, and performance, personality adjustment during sleep deprivation, and the pathology and socialization of sleep.

Professor Murray, an experimental psychologist, is a clear and forceful writer as well as an imaginative investigator. Starting his research on sleep at the Walter Reed Army Institute of Research in the late fifties, he has since kept it aroused at Syracuse. From the time of his first experience with subjects deprived of sleep, he noted the relevance of motivation. Thereupon, he began "a systematic analysis of the sleep motive." The contents of this book show clearly that he has assimilated the many major theories of both motivation and of sleep. This, however, is not a summary of the theories of others. Instead, it is an original amalgam of elements selected here and there—skillfully forged into a very convincing argument—or theory. The "sleep motive," Murray explains, "is analogous to the hunger motive, the sex motive, the curiosity motive, the achievement motive." He is apparently concerned that for sleep, less perhaps than for other motives, there seems to be no particular goal and that the consummatory response is less active. By varying the amount of sleep loss, the sleep hunger or motive can also be varied, as with other drives. Thus sleep seems to fit the typical conception of a motive. Not a habit, not a reflex, not an instinct, not just a consequence of fatigue, it should be examined as a motive. This book therefore examines the phenomenon of sleep in a variety of contexts, in each of which the motive to sleep has the characteristics usually attributed to other motives in that particlular context. Here are a few samples of the types of contextual questions he asks and answers affirmatively.

(1) Does sleep serve some function in biological survival analogous to eating and oether motivated activities? (2) Are there more similarities than differences between the physiological mechanisms underlying this active inhibitory process of sleep and those for other motives? (3) Does the goal act of sleeping like other motives induce learning and performance of special responses, and is it incompatible in performance with the goal acts of other motives? (4) In regard especially to the "energizing" concept of motivation, is the difficulty of arousal systematically related to sleep deprivation? (5) Is the physiological motive of sleep modifiable by personal and social processes, and also does this motive when frustrated carry similar affective, cognitive, or defensive adaptations as do other motives? (6) Does the pathology of sleep support an interpretation of sleep as a motive?

What are the values of such an analysis? Implicit in Murray's argument is the conclusion that meager as it may be, our understanding of other motives outstrips our knowledge of the motives proper to sleep. Still, his purpose as stated in his conclusion "was to add to the list of motives entering the theoretical discourse on the nature of motivation. The relationship between the sleep motive and homeostatic balance provides additional evidence for the idea that motives are based on evolved neural mechanisms partially correlated with bodily needs rather than simple reflexes growing out of tissue deficiencies. . . . The analysis suggests a closer examination of the relationship between motivation and arousal."

One problem which Murray confronts often is the nature of the "activating" properties. Specifically, "the material on learning, performance and arousal suggests that, while the sleep motive has directional properties, it does not function as a general energizer." Finally, he argues for intensive and extensive analysis of all motivational variables. "Perhaps, then, we would be ready for a really meaningful theory of motivation." Murray's thesis on sleep as a motive is likely to generate a great deal of research by those directly concerned with sleep and also by those working with motivation.

Foulkes's book on eye-movement and dream reports appears less likely to stimulate more research. In fact, it need not, for such studies are already at peak quantity. What, however, do we learn from his work? For one thing, we know that if we awaken a sleeper during rapid eye-movements (*REM*s), we increase the probability of a dream-report. This knowledge alone makes the book useful. Secondly, we learn here that *REM*s are cyclical, infrequent, and as essential to the night's sleep activity as are any other of the component parts of the process. Thirdly, we now see that *REM*s are only one of several physiological activities that occur simultaneously with them. Fourthly, we know that specific characteristics of the eye-movement can tell us nothing for certain about the content of the dream that is presumably going on at the time of these eye-movements.

The dreams which by *REM* discovery are now more readily collected are known to be manipulable by experience; and, as Foulkes himself showed some four years ago, the report probability is influenced by the nature of the question. One must ask not what the subject was dreaming, but what he was thinking about, if one wishes to get a report during most non-*REM* periods. The evidence does not support the idea that dreams are produced by unresolved tensions nor do they serve to guard sleep.

Foulkes describes much of the recent work on *EEG*s, *REM*s, and dreaming. His object he states in the preface as follows: "The goal of this book is the study of mental experience during sleep—what it is, and why it is what it is." This little volume tells the reader something about *what* the dream experience may be, and something of timing and causes of reports, but it tells little about "why it is what it is." In the postscript his claims seem less. "If I have any particular hopes for this book, it is that it may have pointed out that the tools are increasingly available for the psychological investigation of sleep." The postscript also compares psychological with physiological research accomplishment concluding that "psychology's immediate past, present, and prospects for the future seem somewhat dim by comparison."

Both books are to be recommended, Foulkes's for an up to date account of the major area of interest in sleep by many psychologists, psychiatrists, and psychoanalysts; namely, *REM*s and dreaming. With so much attention we may expect useful findings, or alternately *REM* studies may suffer the fate, say, of the study of mental content through introspection. In the past, similar clusters of activity have generated a lot of heat and even some curiosity but have produced little if any solid scientific contribution. Murray's, by contrast, boldly and unapologetically throws out a full blown hypothesis well enough stated that it will be likely to stir theoretical debate and experimentation. The Murray book assumes that the reader possesses considerable technical sophistication but Foulkes beams his summary and interpretation to "the interested layman as well as the specialist." Both should be read with profit by the psychologist.

Virginia Military Institute DEAN FOSTER

The Content Analysis of Dreams. By CALVIN S. HALL and ROBERT L. VAN DE CASTLE. New York, Appleton-Century-Crofts, 1966. Pp. xi, 320. $7.00.

The authors, as is well known, are members of the Institute of Dream Research, and the preparation of this book was made possible by a grant from the National Institute of Mental Health of the United States Public Health Service. The authors propose that objective analysis of the manifest content of dreams will provide "one of the most powerful tools for comprehending the dynamics of human behavior." From such a source, and with such support, one looks into the book with considerable anticipation.

Hall's earlier papers on the conceptual theory of dreams (C. S. Hall, A cognitive theory of dreams, *J. gen. Psychol.*, 48, 1953, 169-186) were extremely interesting because they challenged psychoanalytic orthodoxy about dreams. The present work, however, only supports one's suspicion that Hall has taken all that is dynamic out of dreams and put nothing in its place.

In the reviewer's opinion, the book lets one see content analysis at its best, in all its technical aspects—the instructions, norms, and classifications are concise; they are constructive and methodically presented. It is said to be the most comprehensive as yet devised. The book begins with a testimony to the value of content analysis, a method which has many advocates and which has been put to excellent use in almost every humanistic field, in political theory, literature, history, biography, linguistics, folklore, and communication-theory, not to mention every branch of personality and psychotherapy research.

Hall and Van de Castle take the point of view that dreams are like plays on a stage, with settings, characters, and props. Actions occur, which may be classified as aggressive, friendly, sexual, successful and the like; and everything may have attributes such as color, size, density, temperature, velocity, and so on to qualify or modify it. The plays, however, have no plots; and most astonishing, the dreamer is not allowed on this stage—any references to the dreamer, whether the *me*, or the *I*, or the *self*, are not in the plot. The system is said to be empirical, though three scales of a theoretical nature are included: they concern castration anxiety (*CA*), castration wish (*CW*) and penis envy (*PE*); oral activities—incorporation (*OI*), emphasis (*OE*); and regression (*RE*). These concessions to dynamic theory, however, are purely categorical: to dream that one "couldn't get his gun to work" is castration wish, and that one has a sword in one's hand is penis envy, in the manner of dreambooks.

The content is recorded on 5 × 8 in. cards, using letters of the alphabet such as we observe above (*CA*, *CW*, etc.).

In the opening chapter the authors argue, somewhat naïvely it must be said, for quantification of general applicability. Chapter 2 puts together some definitions: in particular, the manifest content of dreams, it is asserted, has "great psychological significance"—an assertion which might have had better support had the material of Chapter 15 been used for that purpose.

Chapter 3 defines the system, and from here to Chapter 13 all is a model of what content analysis as such, in its technical respects, can be. The classes are clearly described, the scoring system adequately illustrated, and all is concise and to the point. Norms, based on 1000 dreams of college students age 19 to 25 yr. are provided in Chapter 14, and these tables are almost all that one could desire. Chapter 15 reviews selected literature on scales of content analysis that have been used by others; Hall and Van de Castle might have done well to start their book with the substance of this chapter, because it provides at least some good reasons for using the manifest content of dreams, and categorical classes in the ancient manner of Aristotle, as basic to a certain kind of study of dreams. There are useful appendices, one listing all the objects and their frequencies, reported in the 1000 dreams. Another gives an example of scoring a dream series.

What, then, can be said of this book? Though the present reviewer is not averse to quantification, nor, does he undervalue content analysis, he is critical of certain of its trends. Indeed, he is of the opinion that some of the most significant contributions in certain fields of study owe everything to the method in empirical respects. What could be more compelling, for example, than the use of the method by Lasswell's school of thought to study the fate of the concept of 'democracy' in international politics over the period 1890 to 1940? What bothers the reviewer is its prodigious use to such little apparent affect.

Consider, for example, the norms. With regard to objects (p. 161), the norms indicate that men report implements more frequently than women, and that women report household articles and clothes more frequently than men. Nothing more significantly different is brought to light! This was also the case in the study by Griffith, Miyago, and Tago (R. M. Griffith, O. Miyago, and A. Tago, The universality of typical dreams: Japanese vs. Americans, *Amer. Anthropol.*, 1958, 60, 1173-1179) in which the only significant differences concerned *finding money*, and *killing someone*. Americans find more money than Japanese; American men report more killings in their dreams than American women, and conversely Japanese women report more killings than Japanese men. These are the only significant differences in a list of 34 items of manifest content of dreams. Hall and Van de Castle refer also to a study by Colby (see page 229) of 752 words as items of dream content. From 400 patients (200 men and 200 women) undergoing psychotherapy and from 400 students (200 men and 200 women), only seven words showed any significant difference (5% level of significance) between patients and students—men referred more to *vehicle, travel, auto,* and *hit,* and women more to *home* and *cry.*

There is no reason, perhaps, why one should expect a sex difference in the manifest content of dreams, but where a significant difference does emerge, it is surely evidence of broad cultural influences rather than of anything particular to dynamic theory of personality.

The outcome of all of this effort is no doubt of interest to someone. What it suggests to this reviewer is massive identities in the manifest content of dreams. No matter how one cuts the population-cake—into sex difference, social class, income-groups, or the like—significant differences in manifest content (when categorized in very broad or universal terms) are likely to be few and far between; and where differences appear they will rarely occasion much surprise. Moreover, we might go a step further and suggest that this will be the same with respect to what people read in the mass media, or look at in television, or what they talk about in everyday conversations. The concern is with the outward form of commonplace language, not in the sense of its style or grammar but in the sense that all "talk" can of course be categorized into settings, characters, props, activities, and attitudes as a matter of merely logical, not necessarily of psychological necessity. This is to say, then, that Hall and Van de Castle have made an empty cardboard model of every dream they analyze, with cardboard settings, cardboard characters, cardboard props, and all else—all drained completely of psychological and dynamic significances.

I am not sure how far I should take the above line of thought, because I am still puzzled about much else in Hall and Van de Castle's book. They altogether miss George's point about the uniqueness of some events in behavior (Chapter 1); and it is difficult to accept their assertion (page 21) that it is doubtful how far dreams really tap deeper levels of personality dynamics. To say (page 26) that in content analysis the system of classification "should grow out of the material and reflect its intrinsic nature" requires clarification as to what is to be regarded as intrinsic and what not. What is intrinsic in Hall and Van de Castle's case is apparently the Aristotelian categorization of things, persons, and events. The outcome, surely, speaks for itself. The truth is that content analysis in science can only be a test of synthetic propositions, that is, of a body of theory of some kind.

What, then, can be theoretical about dreams? It was to this question that I thought Hall was addressing himself in his earlier papers. For Hall is right, in my opinion, to challenge the Freudian standpoint that dreams are necessarily regressive, in the sense that, being fantasy, they substitute for painful emotions. The basic paper is Freud's own, on the "Relation of the Poet to Daydreaming." The contrast is between the real world of painful emotions and the "playworld" of daydreams, dreams, and perhaps of much else that we think of as real. For Freud, happy people have no need of fantasies—only the disatisfied do. This I find impossible to accept. Rather, when Pope wrote his "The Rape of the Lock," as when the Fuller Brush salesman of Hall and Van de Castle's book dreams of lecturing to a class of 30 eighteen-year old girls about the superior merits of hair brushes over nylon brushes, the fantasies may be pure fun, pure play, pure delight, pure serenity, pure self-striving and enjoyment. The days are surely over when concepts of pleasure can be left, theoretically, at the non-self, id-process levels of classical psychoanalysis. Dreams need a new looking-at in communication-pleasure terms, in relation to genuine theories of play, not cardboard models of a stage, its settings, characters, and props. It is something of the kind I had expected from Hall following his provocative early papers. Might he not be encouraged to study dreams, by content analysis if need be, in genuine theoretical and not merely Aristotelian terms?

University of Missouri WILLIAM STEPHENSON

Eye and Brain: The Psychology of Seeing. By R. L. GREGORY. New York, McGraw-Hill, 1966. Pp. 251. $4.95.

Gregory's book on visual perception faces a dilemma common to books in this area. Basically, the summary and simplification necessary in a book aimed at the non-specialist are difficult simply because there are too many fundamental facts still missing from the field for safe theoretical generalizations. The author of a popular book on perception must therefore steer between the Scylla of presenting a series of interesting but unrelated phenomena and the Charybdis of making unwarranted assumptions about their interpretation and meaning. Probably no book for the general reader can be written which avoids one pitfall without risking the other. What course has Gregory navigated?

Gregory has avoided writing a book which would be simply a collection of visual demonstrations, uninterpreted and unconnected. The book, well written and strikingly illustrated, offers a definite point of view. It begins by introducing the reader to the pre-conditions of perception—the fundamental facts of the physics of light, the physiology and anatomy of the eye and the visual pathways, and the basic sensory relationships. A commendable feature of the book is that both psychological and physiological findings are presented in discussing the sensory functions of the eye. For Gregory, perception, however, is much more than the registering of patterns of stimulation. A percept is "a hypothesis suggested and tested by sensory data" (p. 12). A hypothesis-testing model of perception within the limitations established by the sensory systems serves as the theoretical orientation of the book. Much of the attraction of this book to the general reader resides in the varied application of this view. Hypothesis-testing in perception is related to the illusions to be experienced by astronauts in space where the visual system would lack essential information for veridical perception; reversible figures, such as the Necker cube, which give no clue as to their correct interpretation; impossible figures, which give contradictory information, and therefore allow no unique interpretation; and to the problems of the artist who cannot hope to provide on a canvas all the cues for depth present in reality.

The book relates perceptual phenomena to specific mechanisms as well as providing a general functional interpretation. The phenomena of the perception of movement are discussed in terms of the ways motion is signalled by the shift of images across the retinas and the rotation of the eyes in the head. These two signal systems provide the basis for explaining why the world remains fixed when we move our eyes, the autokinetic illusion, the waterfall illusion, and the phi phenomenon. Similarly, the Young-Helmholtz theory of color is used to integrate the facts about color vision and color blindness, and the visual illusions of size and direction are explained as resulting from the habits we have acquired in learning to perceive space accurately. In suggesting mechanisms for integrating diverse perceptual phenomena, Gregory is aware of the complexity of the problems and that the limitations of the models proposed are soon reached. He points out that the image-retina and eye-head mechanisms can not account for the illusory perceptions of motion experienced when the perceived position of a target is changed relative to a point of reference, and that the basic mechanisms of color vision can not account for the attributive and functional characteristics of surface colors.

In avoiding the dangers of unconnected illustrations, Gregory unfortunately presents many of his own views as if they were established facts. Though the illusions have been extensively studied, there is still no generally accepted understanding of what causes them. There is considerable evidence that the patterned illusions are affected by both judgmental factors and by physiological interactions in the visual pathways. Gregory, however, ascribes these illusions primarily to judgment, that is, he suggests that as a result of the habits we have acquired in order to achieve constancy, the parts of illusory figures which represent distant objects are enlarged and those representing near objects are diminished. In like manner, the autokinetic illusion is ascribed to abnormal signals in the eye-head system controlling the movement of the eyes. Perhaps Gregory is correct on both counts, but there is yet no conclusive evidence for his explanation. The book also unfortunately attempts to argue from demonstration. For example, from the demonstration that metameric matches of yellow (monochromatic yellow and a mixture of red and green) continue to match in different states of adaptation, the author argues that there is no specialized yellow receptor. The most that can be concluded, however, is that the assumption that there are only three kinds of color receptors is consistent with the fact that colors which match in one state of adaptation continue to match in other states of adaptation. Though there is evidence for asserting that the cue-properties of stimuli as well as the relationships affect color perception, many would also question the statement that Land's demonstration "brings out the complicated additions made by the brain to sensory information when sensations are organized into perceptions of objects" (p. 125).

Gregory's emphasis on functionalism adequately accounts for the empirical facts showing that perceptions are affected by attitudes, expectancies, and motives. The book, however, does not adequately deal with the fact that perceptions are also typically structured. Perceptual processes involve the integration and organization of patterns of stimulation as well as hypothesis-testing. Gregory assumes that the cue properties of stimuli are sufficient to define the integration of a stimulus-pattern but does not present the problems posed by this view. For example, in discussing the Ames distorted room, he asserts that we are so used to rectangular rooms that rather than perceive a distorted room we perceive distortions in the sizes of people and objects as they are moved from one side of the room to the other. In terms of past experience it would, however, seem more likely that an observer should see the room as distorted rather than see people and objects grow and shrink as they move about the room. What he fails to consider is that the perception of size is determined not solely in terms of a "best bet" but also in terms of the relationship of an object to a frame of reference. In general, the book does not adequately deal with the organizational processes in perception. Aside from describing induced movement ("the brain takes the best bet and tends to accept that movement is of smaller rather than larger objects" p. 114), the problems posed by the configurational factors in movement perception are not discussed. Nor does Gregory include a discussion of the principles of organization that have emerged from the study of visual patterns or the recent attempts to formulate these principles in terms of concepts of an economical or simple encoding.

In summary, this book provides an exceptionally good introduction to a variety of perceptual problems. It, however, often expresses a personal point of view which may arouse the protests of knowledgeable people.

University of Oregon JACOB BECK

Unobtrusive Measures: Nonreactive Research in the Social Sciences. By EUGENE J. WEBB, DONALD T. CAMPBELL, RICHARD D. SCHWARTZ, and LEE SECHREST. Chicago: Rand McNally & Company, 1966. Pp. xii, 225. $3.50.

The purpose of this work is to broaden the methodological outlook of social scientists. Citing the near-monopoly currently enjoyed by interviews and questionaries, the authors direct attention to approaches which "do not require a respondent and do not themselves contaminate the response" (p. 2). The book opens in a humorous spirit; the Preface informs the reader that the manuscript had circulated under such titles as: *Oddball Measures* or *The Bullfighter's Beard.* To the reviewer it seems entirely appropriate that a call to divergent methodological thinking should be irreverently expressed. Perhaps, if sober argument fails, levity may succeed in loosening stiffened research muscles.

Contents of the volume extend in two general directions: to principles of scientific measurement and to a compendium of measurement possibilities; orientations developed in the first effort provide a background for the second. With regard to the logic of measurement, the flavor of the message is expressed in the title to Chapter I: "Approximations to Knowledge." The argument given here does not follow the usual insistence about selecting and sharpening the best measure for a variable; instead the plea is for multiple and overlapping measurement. Measurements which overlap serve as checks upon one another as well as provide fullness and detail to the variables they represent. Knowledge, it is contended, is approximate and is gained by data "triangulation."

Although the book is occasionally light in style, the authors are neither naïve nor flippant about the threats to valid inference posed by various research methods. The discussion of methodological pitfalls is short but adequate. Description of distorting effects of reactive methods is knowledgeable; but, it turns out, any single method is prone to some kind of error or misrepresentation. What is needed is measurement of various kinds at a number of points. The idea is expressed via a geological model:

> Any given theory has innumerable implications and makes innumerable predictions which are unaccessible to available measures at any given time. The testing of the theory can only be done at the available outcroppings, those points where theoretical predictions and available instrumentation meet. Any one such outcropping is equivocal, and all types available should be checked (p. 28).

After the introductory discussion, the authors devote five chapters to illustrations and evaluations of nonreactive measurements. Physical evidence is first. Behavior leaves traces in the form of erosions and accretions. With proper supplementary knowledge, sometimes with controls, it is possible to infer the extent to which certain behaviors have occurred. For example, the popularity among children of certain museum exhibits can be assessed by counting noseprints on the display window.

Next to be described are examples of archival and other documentary sources of behavioral information. These range from bureaucratic records to private letters and even suicide notes. Methods for checking and correcting the biases that exist in such sources are discussed. The principle of overlapping data applies to correcting measurements as well as to checking hypotheses. For example, one might hope to study typical family size at a point in the past by using popular fiction. If data on family size, as given by reliable census sources, can be matched with a part of data obtained from fiction, one can estimate the accuracy of the fictional data; if this comparison at the overlap is favorable, the fictional data of family-size become available for places or periods which lack census data. Information from the comparison can also be employed to create corrections or compensations in the fictional data.

The last set of methods is observational. A chapter is devoted to simple and naturalistic approaches to behavioral observation and recording; another to the use of hardware and employment of contrived or provocative situations. The authors point out that the latter measures can elicit kinds and amounts of data which simple observation yields grudgingly, if at all. Simple observation is also presumedly handicapped by the reactive effects produced when subjects are aware of observation. The contention seems to be that the only good, *i.e.* nonreactive, observation is surreptitious observation. To the reviewer this is a portentous conclusion and should not be accepted without close examination. Once investigators are convinced of the requirement of subject unawareness, the utilization of subject-deceiving techniques (the hidden microphone or camera) seems almost mandatory; these, in turn, raise serious issues regarding ethical and public relations. The authors disclaim sanction of deceitful practices but must describe them all the same. The discussion inevitably supports the use of concealed methods if it strongly states the reactive effects inherent in direct and open observation depends upon his sensitivities, his interpretation of the observation's purpose, the compellingness of other aspects of his situation, the length of time he has been under observation, and other factors. In terms of the alternatives, the field cannot afford simply to *assume* that open observation produces significant reactive effects.

The authors conclude with a return to issues of research inference. As psychology has become more sophisticated and more rigorous, the variety of problems and methods has correspondingly lessened. To widen our understanding of phenomena we need to forego the escape offered by definition in terms of one or two operations. We need to employ a variety of measures which, although prone to error, are erroneous in different fashions than are interviews and questionaries. Much of the conclusion is a helpful recapitulation of issues and solutions previously discussed.

In assessing *Unobtrusive Measures* one can claim that some methodological points are hardly new. True, but these points are neatly made; in fact, there are a number of highly instructive and quotable sections in this little volume. A more substantial complaint can be that the principles of research measurement are insufficiently worked out in the chapters on method. The book does not demonstrate thoroughly how overlapping data can be collected and triangulated in pursuit of a significant research problem. Methods are described and evaluated but not truly integrated into the stimulating "outcropping" orientation developed in the introduction. Perhaps this more disciplined work is reserved for a future monograph.

As it stands this book contains much more to appreciate than to criticize. It offers a fund of ideas and examples which can keep the open-minded investigator engaged for some time. The presentation can enrich methodological repertoires and freshen research outlooks.

University of Kansas

PAUL GUMP

Outlines of a Theory of the Light Sense. By EWALD HERING. Translated by LEO M. HURVICH and DOROTHEA JAMIESON. Harvard University Press, 1964. Pp. xxvii, 317. $9.95.

A translation of Hering's classic but still pertinent work on color has long been needed. The Hurviches, who are unquestionably the most competent interpreters of Hering in the English-speaking world, have given us an excellent translation, have written an informative introduction, and have compiled a useful list of references to primary sources. The translation is from the *Grundzüge der Lehre vom Lichtsinn* (1920), a posthumous volume which embodies the results of more than forty years of reseach and reflection on the problems of color perception. This is Hering's most important book. It should be in every academic library and in the personal library of every psychologist interested in perception.

The sad thing about Hering is that in the textbooks he is always matched against Helmholtz; and Helmholtz, having been translated early and well, has had the better press. Helmholtz and Hering had rival theories of color, each of which in its own way has been fruitful. Helmholtz has been tagged as the empiricist, Hering as the nativist, a distinction which has often led to misunderstandings. Both were rigorously empirical scientists, willing to look to past experience for explanations where this seemed warranted by the evidence. They differed, however, in their interests. Helmholtz was first and foremost a physicist, Hering a physiologist and psychologist. They looked at the same phenomena, at the constancy of color for example, but, whereas Helmholtz disposed of the constancy problem by invoking unobservable processes of unconscious inference, Hering sought for an explanation in terms of the built-in machinery of the body: pupillary adjustment, retinal adaptation, and contrast, with past experience playing a supportive role in the form of memory color. It was Hering, not Helmholtz, who revealed the constancies as a richly rewarding field for empirical research.

Hering's book is a landmark in the history of the psychology of perception. It was Hering who inspired David Katz's studies of color and touch and, less directly, the Gestalt studies of perceptual organization; he was without any question the best of the early experimental phenomenologists. The book would be worth reading, however, even if it were now appearing for the first time, for it represents a point of view which is just as important in the modern age of the computer as it was at the end of the last century.To the psychologist who is allowing the computer to decide what is and what is not a legitimate psychological datum Hering would say, as he did to the physicalist Helmholtz: Look first at the phenomena before you build your model; and if your model cannot do justice to all the phenomena, don't simply discard the phenomena which don't fit; better to scrap the model and try again. If Hering were alive today he would probably be fascinated by the possibilities of the computer, but he would not permit the computer to dictate his science.

Cornell University ROBERT B. MACLEOD

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review).

ALTMAN, JOSEPH, *Organic Foundations of Animal Behavior.* New York, Holt, Rinehart & Winston, 1966. Pp. xiv, 530. $12.95.

AMERICAN BEHAVIORAL SCIENTIST, *Recent Publications in the Social and Behavioral Sciences, The ABS Guide. 1966 Supplement.* Berkeley, Calif., Sage Publications, 1966, Pp. 220. $12.50.

ASTRUP, CHRISTIAN, and NOREIK, KIELL, *Functional Psychoses. Diagnostic and Prognostic Models.* Springfield, Ill., C. C Thomas, 1966. Pp. xix, 176. $9.50.

ATTESLANDER, PETER, and GIROD, ROGER, *Soziologische Arbeiten, Travaux sociologiques. Sociological Contributions.* Bern and Stuttgart, Verlag Hans Huber, 1966. Pp. 348. Fr. 28—.

BAKAN, DAVID, *The Duality of Human Existence. An Essay on Psychology and Religion.* Chicago, Ill., Rand McNally, 1966. Pp. ix, 242. $6.50.

BALBERNIE, RICHARD, *Residential Work with Children.* Long Island City, N.Y. Pergamon Press, 1966. Pp. x, 244. $7.25.

BALTZELL, E. DIGBY, *Philadelphia Gentlemen. The Making of a National Upper Class.* New York, Free Press, 1966 (First published 1958). Pp. ix, 440. $2.95 (paperback).

BARROW, ROSCOE L., and FABING, HOWARD D., *Epilepsy and the Law. A Report on Legal Reform in the Light of Medical Progress.* Rev. 2nd ed. Pp. xvi, 174. $5.95 (paperback).

BEALS, ALAN R., and SIEGEL, BERNARD, *Divisiveness and Social Conflict. An Anthropological Approach.* Stanford, Calif., Stanford Univ. Press, 1966. Pp. x, 185. $6.00.

BIDDLE, BRUCE J., and THOMAS, EDWIN J. (eds.), *Role Theory. Concepts and Research.* New York, Wiley, 1966. Pp. xiv, 453. $8.95.

BIGGAR, JEAN, *Psychotherapy and Child Development. An Introduction for Students.* New York, Barnes & Noble, 1966. Pp. xi, 142, $4.50.

BIRCH, DAVID, and VEROFF, JOSEPH, *Motivation: A Study of Action.* Belmont, Calif., Brooks/Cole Publ. Co., 1966. Pp. ix, 98. $1.50 (paperback).

BRECHER, RUTH and EDWARD (eds.), *An Analysis of Human Sexual Response.* New York, New American Library, 1966. Pp. xiv, 318. 75¢ (paperback).

BROCKBANK, REED, and WESTBY-GIBSON, DOROTHY (eds.), *Mental Health in a Changing Community.* New York and London, Grune & Stratton, 1966. Pp. ix, 163. $6.50.

BRUNER, JEROME, S., OLIVER, ROSE R., GREENFIELD, PATRICIA M., *et al., Studies in Cognitive Growth.* New York, Wiley, 1966. Pp. xviii, 343. $7.95.

CAMPBELL, ANGUS; CONVERSE, PHILIP E.; MILLER, WARREN E.; and STOKES, DONALD E., *Elections and the Political Order.* New York, Wiley, 1966. Pp. ix, 385. $8.75.

CHERRY, COLIN, *On Human Communication. A Review, A Survey and a Criticism.* 2nd ed. Cambridge, Mass. and London, Eng., M.I.T. Press, 1966. Pp. xiv, 337. $10.00.

CHURCH, JOSEPH, *Language and the Discovery of Reality.* New York, Vintage Books, 1966 (first published 1961). Pp. xv, 245. $1.95 (paperback).

CLARKE, ANN M., and CLARKE, A. D. B. (eds.), *Mental Deficiency. The Changing Outlook.* Rev. ed. New York, Free Press, 1966. Pp. xxi, 596 $10.95.

COMMITTEE ON THE COLLEGE STUDENT, GROUP FOR THE ADVANCEMENT OF PSYCHIATRY, *Sex and the College Student.* Greenwich, Conn., Fawcett Publications, 1966. Pp. 158. 75¢ (paperback).

DANA, RICHARD H., *Foundations of Clinical Psychology. Problems in Personality and Adjustment.* Princeton, N.J., D. VanNostrand, 1966. Pp. viii, 322.

DAY, R. H., *Perception. Introduction to General Psychology: A Self-Selection Textbook.* Dubuque, Iowa, W. C. Brown Co., 1966. Pp. 46. 65¢ (paperback).

De Levita, David J., *The Concept of Identity*. New York, Basic Books, 1967. Pp. vii, 209. $6.00.

Diamond, Robert M., and Woodward, John C., *The Amateur Psychologist's Dictionary*. New York, ARC Books, 1966. Pp. 159. $2.50, cloth; .95 paper.

Dickson, William J., and Roethlisberger, F. J., *Counseling in an Organization. A Sequel to the Hawthorne Researches*. Boston, Mass., Grad. School of Bus. Admin., Harvard Univ., 1966. Pp. xviii, 480. $10.00.

Diggory, James C., *Self-Evaluation. Concepts and Studies*. New York, Wiley, 1966. Pp. xiii, 477.

Doob, Leonard W., *Communication in Africa. A Search for Boundaries*. New Haven, Conn., Yale Univ. Press, 1966 (first published 1961). Pp. xvi, 406. Cloth, $8.50; Paper, $2.45.

Downes, David M., *The Delinquent Solution. A Study in Subcultural Theory*. New York, Free Press, 1966. Pp. xi, 284. $6.95.

Ehrenwald, Jan, *Psychotherapy: Myth and Method. An Integrative Approach*. New York, Grune & Stratton, 1966. Pp. x, 212. $7.75.

Eissler, K. R., *Sigmund Freud und die Wiener Universität. Über die Pseudo-Wissenschaftlichkeit der jüngsten Wiener Freud-Biographik*. Bern and Stuttgart, Verlag Hans Huber, 1966. Pp. 191. Fr. 25—. (paperback).

Evans, C. R., and Robertson, A. D. J. (eds.), *Brain Physiology and Psychology*. Berkeley, Calif., Univ. of Calif. Press, 1966. Pp. viii, 280. $5.00 (paperback).

Evans, Jean, *Three Men. An Experiment in the Biography of Emotion*. New York, Vintage Books, 1966 (first published 1950), Pp. xviii, 297. $1.65 (paperback).

Faulkner, E. J. (ed.), *Man's Quest for Security. A Symposium*. Lincoln, Nebr., Univ. of Nebraska Press, 1966. Pp. xiv, 197. $5.00.

Fernald, L. Dodge, and Fernald, Peter S., *Overview of General Psychology. A Basic Program*. Boston, Mass., Houghton Mifflin, 1966. Pp. xi, 354. $3.50 (paperback).

Fogel, Lawrence J., Owens, Alvin, and Walsh, Michael J., *Artificial Intelligence through Simulated Evolution*. New York, Wiley, 1966. Pp. xii, 170. $9.95.

Fraisse, Paul, *Praktikum der Experimentellen Psychologie*. Bern and Stuttgart, Verlag Hans Huber, 1966, Pp. xvi, 379. Fr./DM 38.—

Gaito, John, *Molecular Psychobiology. A Chemical Approach to Learning and Other Behavior*. Springfield, Ill., C. C Thomas, 1966. Pp. xix, 259. $9.50.

Galanter, Eugene, *Textbook of Elementary Psychology*. San Francisco, Calif., Holden-Day, 1966. Pp. xiii, 419. $7.95.

Goffman, Erving, *Behavior in Public Places. Notes on the Social Organization of Gatherings*. New York, Free Press, 1966 (first published 1963). Pp. viii, 248. $2.45 (paperback).

Goldstein, Arnold P., and Dean, Sanford J., *The Investigation of Psychotherapy. Commentaries and Readings*. New York, Wiley, 1966. Pp. xiii, 443. $9.95.

Grob, Gerald, N., *The State and the Mentally Ill. A History of Worcester State Hospital in Massachusetts 1830-1920*. Chapel Hill, N.C., Univ. of North Carolina Press, 1966. Pp. xv, 399. $7.50.

Haerem, Alf T., *A Handbook for Neurotics*. Boston, Mass., Christopher Publ. House, 1966 (first published 1960). Pp. 137. $3.95.

Hall, Calvin, S., *The Meaning of Dreams*. New York, McGraw Hill, 1966 (first published 1953). Pp. xxv, 244. $2.45 (paperback).

Harvey, O. J. (ed.), *Experience, Structure and Adaptability*. New York, Springer, 1966. Pp. viii, 406. $9.00.

Hebron, Miriam E., *Motivated Learning. A Developmental Study from Birth to the Senium*. New York, Barnes & Noble, 1966. Pp. 264. $6.25.

Heckel, Robert V., and Peacock, Lelon J., *Textbook of General Psychology*. Saint Louis, Mo., C. V. Mosby, 1966. Pp. xii, 480. $8.50.

Hellmuth, Jerome (ed.), *Learning Disorders, Vol. 2*. Seattle, Wash., Special Child Publications, 1966. Pp. 423. $10.00.

Hempel, Carl G., *Philosophy of Natural Science*. Englewood Cliffs, N.J., Prentice-Hall, 1966. Pp. x, 116. $1.75 (paperback).

Hiltmann, H., *Kompendium der psychodiagnostischen Tests*. Bern and Stuttgart, Verlag Hans Huber, 1966. Pp. 304. Fr./DM 22.—(paperback).

Hovland, Carl I. (ed.), *The Order of Presentation in Persuasion.* New Haven and London, Yale Univ. Press, 1966 (first published 1957). Pp. x, 192. Cloth, $5.00; Paper, $1.45.

Hudson, Liam, *Contrary Imaginations. A Psychological Study of the Young Student.* New York, Schocken Books, 1966. Pp. 181. $4.95.

Huessy, Hans R., *Mental Health with Limited Resources. Yankee Ingenuity in Low-Cost Programs.* New York and London, Grune & Stratton, 1966. Pp. xii, 124.

Hurlock, Elizabeth B., *Adolescent Development.* 3rd ed. New York, McGraw-Hill, 1967. Pp. xi, 719. $8.95.

Johnson, E. C., *Applied Psychology. Science of Living.* New York, Carlton Press, 1966. Pp. 296. $6.50.

Kaplan, Michael (ed.), *Essential Works of Pavlov.* New York, Bantam Books, 1966. Pp. ix, 392. $1.25 (paperback).

Klineberg, Otto, *Die menschliche Dimension in den internationalen Beziehungen.* Bern and Stuttgart, Verlag Hans Huber, 1966. Pp. 168. Fr/DM 22.

Koch, Helen L., *Twins and Twin Relations.* Chicago, Ill., Univ. of Chicago Press, 1966. Pp. xvi, 302. $6.95.

Köhler, Wolfgang, *The Place of Value in a World of Facts.* New York and Toronto, New American Library, 1966 (first published 1938). Pp. xi, 320. 95¢ (paperback).

Kraines, S. H., and Thetford, E. S., *Managing Your Mind.* Greenwich, Conn., Fawcett Publications, 1966 (first published 1943). Pp. 314. 75¢ (paperback).

Küppers, Waltraut, *Zur Psychologie des Geschichtsunterrichts.* Bern and Stuttgart, Hans Huber and Ernest Klett, 1966. Pp. 159. Fr. 19.80.

Lehtovaara, Arvo, *The Predictiveness of the Finnish Matriculation Examination. A Differential Psychological Study.* Helsinki, Suomalainen Tiedeakatema, 1966. Pp. 288.

Lenneberg, Eric H. (ed.), *New Directions in the Study of Language.* Cambridge, Mass., M.I.T. Press, 1966 (first published 1964). Pp. ix, 194. $1.95 (paperback).

Levine, David (ed.), *Nebraska Symposium on Motivation, 1966.* Lincoln, Nebr., Univ. of Nebraska Press, 1966. Pp. ix, 209. Cloth, $5.95; Paper, $2.75.

Lorr, Maurice (ed.), *Explorations in Typing Psychotics.* Long Island City, N.Y., Pergamon Press, 1966. Pp. viii, 241. $9.25.

Lowen, Alexander, *The Betrayal of the Body.* New York, Macmillan, 1967. Pp. 307. $6.95.

Maccoby, Eleanor E. (ed.), *The Development of Sex Differences.* Stanford, Calif., Stanford Univ. Press, 1966. Pp. 351. $8.50.

MacLean, Isabella C., *Child Guidance and the School. Modern Teaching.* New York, Barnes & Noble, 1966. Pp. 117. Cloth, $3.25; Paper, $1.75.

Manheim, Leonard and Eleanor (eds.), *Hidden Patterns. Studies in Psychoanalytic Literary Criticism.* New York, Macmillan, 1966. Pp. 310. $6.95.

Master, R. E. L., *Sex-Driven People. An Autobiographical Approach to the Problem of Sex-Dominated Personality.* Los Angeles, Calif., Sherbourne Press, 1966. Pp. 287. $6.50.

McGuigan, F. J., *Thinking: Studies of Covert Language Processes.* New York, Appleton-Century-Crofts, 1966. Pp. x, 295. $2.95 (paperback).

McLaren, Anne (ed)., *Advances in Reproductive Physiology, Vol. 1.* New York and London, Academic Press, 1966. Pp. viii, 295. $12.50.

Megargee, Edwin I. (ed.), *Research in Clinical Assessment.* New York, Harper & Row, 1966. Pp. xv, 702. $14.75.

Misiak, Henryk, and Sexton, Virginia Staudt, *History of Psychology. An Overview.* New York, Grune & Stratton, 1966. Pp. x, 499. $12.50.

Moriarty, Alice E., *Constancy and IQ Change. A Clinical View of Relationships Between Tested Intelligence and Personality.* Springfield, Ill. C. C Thomas, 1966. Pp. vii, 223.

Morris, Ramona and Desmond, *Men and Apes.* New York, McGraw-Hill, 1966. Pp. viii, 271. $7.95.

Mullan, Hugh, and Sangiuliano, Iris, *Alcoholism. Group Psychotherapy and Rehabilitation.* Springfield, Ill., C. C Thomas, 1966. Pp. xxi, 325. $12.00.

National Commission on Community Health Services, *Health is a Community Affair.* Cambridge, Mass., Harvard Univ. Press, 1966. Pp. xiv, 252. $5.00.

Page, James D. (ed.), *Approaches to Psychopathology.* New York and London, Columbia Univ. Press, 1966. Pp. xiii, 304. $7.50.

Patterson, C. H., *Theories of Counseling and Psychotherapy.* New York, Harper & Row, 1966. Pp. ix, 518. $9.75.

Pei, Mario (ed.), *Language of the Specialists. A Communications Guide to Twenty Different Fields.* New York, Funk & Wagnalls, 1966. Pp. xii, 388. $7.50.

Pikas, Anatol, *Abstraction and Concept Formation. An Interpretative Investigation into a Group of Psychological Frames of Reference.* Cambridge, Mass., Harvard Univ. Press, 1966. Pp. xiii, 303. $7.00.

Post, Felix, *Persistent Persecutory States of the Elderly.* Long Island City, N.Y., Pergamon Press, 1966. Pp. ix, 100. $4.00.

Redl, Fritz, *When We Deal with Children. Selected Writings.* New York, Free Press, 1966. Pp. xii, 511. $8.95.

Roberts, Leigh M., Halleck, Seymour L., and Loeb, Martin B. (eds.), *Community Psychiatry.* Madison, Wisc., Univ. of Wisconsin Press, 1966. Pp. x, 252. $6.00.

Rock, Irvin, *The Nature of Perceptual Adaptation.* New York, Basic Books, 1966. Pp. x, 289. $8.50.

Rosenberg, Milton J., Hovland, Carl I., McGuire, William J., Abelson, Robert P., and Brehm, Jack W., *Attitude Organization and Change. An Analysis of Consistency Among Attitude Components.* New Haven and London, Yale Univ. Press, 1966 (first published 1960). Pp. x, 239. Cloth, $6.00; Paper, $1.75.

Rosenfels, Paul, *Love and Power. The Psychology of Interpersonal Creativity.* New York, Libra Publishers, 1966. Pp. 232. $4.95.

Rosenthal, Robert, *Experimenter Effects in Behavioral Research.* New York, Appleton-Century-Crofts, 1966. Pp. xiii, 464. $7.50.

Runes, Dagobert D., *Treasury of Thought. Provocative Observations on the Great Topics that Have Challenged Man's Thought Throughout the Ages.* New York, Philosophical Library, 1967. Pp. 395. $6.00.

Sahakian, William S., and Sahakian, Mabel Lewis, *Ideas of the Great Philosophers.* New York, Barnes & Noble, 1966. Pp. viii, 181. $1.50 (paperback).

Savage, R. D. (ed.), *Readings in Clinical Psychology.* Long Island City, N.Y., Pergamon Press, 1966. Pp. x, 811. $23.00.

Schachtel, Ernest G., *Experiential Foundations of Rorschach's Test.* New York, Basic Books, 1966. Pp. vi, 342. $7.95.

Schoeck, Helmut, *Der Neid. Eine Theorie der Gesellschaft.* Freiburg/München, Verlag Karl Alber, 1966. Pp. 432.

Schroder, Harold M., Driver, Michael, J., and Streufert, Siegfried, *Human Information Processing. Individuals and Groups Functioning in Complex Situations.* New York, Holt, Rinehart and Winston, 1967. Pp. viii, 224. $7.95.

Seed, Phillip, *The Psychological Problem of Disarmament.* London, Housmans, 1966. Pp. 74. $1.00 (paperback).

Sherif, Muzafer, *In Common Predicament. Social Psychology of Intergroup Conflict and Cooperation.* Boston, Mass., Houghton Mifflin, 1966. Pp. xv, 192. $4.95.

Shumaker, Wayne, *Literature and the Irrational. A Study of Anthropological Backgrounds.* New York, Washington Square Press, 1966 (first published 1960). Pp. xix, 279. 90¢ (paperback).

Smith, Frank, and Miller, George A. (eds.), *The Genesis of Language. A Psycholinguistic Approach.* Cambridge, Mass., M.I.T. Press, 1966. Pp. xii, 400. $10.00.

Smith, Wendell I., and Moore, J. William, *Conditioning and Instrumental Learning. A Program for Self-Instruction.* New York, McGraw-Hill, 1966. Pp. xv, 169. Cloth, $4.95; Paper, $2.50.

Spielberger, Charles D. (ed.), *Anxiety and Behavior.* New York and London, Academic Press, 1966. Pp. xiv, 414. 78s.

Sternbach, Richard A., *Principles of Psychophysiology. An Introductory Text and Readings.* New York and London, Academic Press, 1966. Pp. xiii, 297. $7.50.

Stollak, Gary E., Guerney, Bernard G., and Rothberg, Meyer (eds.), *Psychotherapy Research. Selected Readings.* Chicago, Ill., Rand McNally, 1966. Pp. vi, 706. $6.50.

STONES, E., *An Introduction to Educational Psychology.* New York, Barnes & Noble, 1966. Pp. 424. Cloth, $6.50; Paper, $3.50.

TRICE, HARRISON M., *Alcoholism in America.* New York, McGraw-Hill, 1966. Pp. xi, 152. Cloth, $3.95; Paper, $1.95.

VANDERPLAS, JAMES M., *Controversial Issues in Psychology.* Boston, Mass., Houghton Mifflin, 1966. Pp. x, 438. $7.50.

VON ZEPPELIN, ILKA S., *Die Variablen der Holtzman Inkblot Technique in ihrer Beziehung zur Introversion und Extraversion.* Bern and Stuttgart, Verlag Hans Huber, 1966. Pp. 102. Fr. 14.80 (paperback).

WARNER, SAMUEL J., *Self-Realization and Self-Defeat.* New York, Grove Press, 1966. Pp. 215. $6.00.

WEINER, IRVING B., *Psychodiagnosis in Schizophrenia.* New York, Wiley, 1966. Pp. xiv, 573.

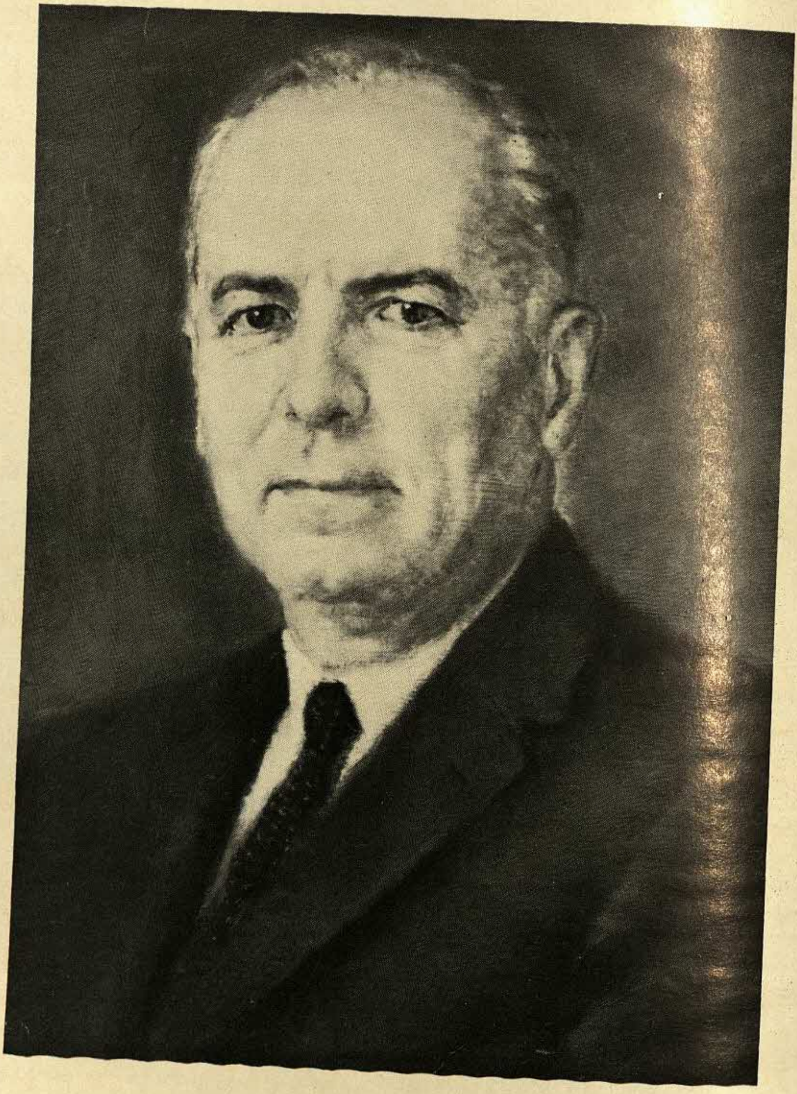
WERTHAM, FREDRIC, *A Sign for Cain. An Exploration of Human Violence.* New York, Macmillan, 1966. Pp. 391. $6.95.

WHIPPLE, DOROTHY V., *Dynamics of Development: Euthenic Pediatrics.* New York, McGraw-Hill, 1966. Pp. xvi, 648. $15.00.

WHITE, ROBERT W., *Lives in Progress.* 2nd ed. New York, Holt, Rinehart and Winston, 1966. Pp. x, 422. Cloth, $6.95; Paper, $4.95.

WOLMAN, BENJAMIN B., *Vectoriasis Praecox or the Group of Schizophrenias.* Springfield, Ill., C. C Thomas, 1966. Pp. xi, 371. $12.50.

WOLPE, JOSEPH, and LAZARUS, ARNOLD A., *Behavior Therapy Techniques. A Guide to the Treatment of Neuroses.* Long Island City, N.Y., Pergamon Press, 1966. Pp. ix, 198. $2.75 (paperback).



Kenneth W. Spence

(See page 314)

THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. STANLEY HALL

VOL. LXXX JUNE, 1967 NO. 2

TIME-ERROR IN COMPARISONS OF PREFERENCES FOR MUSICAL EXCERPTS

By SOON DUK KOH, Hastings State Hospital, Minnesota

The judgmental bias due to the order of presentation of stimuli, or the constant deviation of the point of subjective equality from the standard stimulus, is traditionally called the time-error (*TE*). Studies of this phenomenon in psychophysics have clarified many underlying features of comparative judgment and significantly contributed to the derivation of judgmental theories.[1] Although the similarity between affective and psychophysical judgments has been well demonstrated since Fechner's experimental esthetics,[2] the attempt to clarify the nature of affective judgments by means of the *TE* has been less fruitful.

Beebe-Center and his associates presented two simple tones or odors successively to *S*s and found a tendency for the second stimulus to be preferred (negative *TE*).[3] Choo, however, found no consistent *TE* with tones and rhythms, but found predominantly negative *TE* with melodies.[4] Brighouse and Koh also found inconsistent *TE* in successive comparisons of colors and achromatic designs, but predominantly negative *TE* with reproduction of paintings.[5] The main issue addressed in the present study, therefore, is whether the *TE* can be reliably and consistently observed in affective or pleasantness-unpleasantness comparisons, and if so, under what conditions.

A review of the studies on time-error (*TE*) in psychophysics indicates that it is a function of variables such as the stimulus-level, stimulus-range, background, and time-interval between stimuli.[6] The inconsistency observed in the studies of affective *TE* may, therefore, have resulted from inadequate control of these variables. This series of studies is, therefore, designed to investigate (1) the effect of stimulus-level and range, and (2) the effect of a 'background' stimulus, respectively, on appearance of the *TE* in affective judgments comparing *S*'s preference for musical selections.

---

* Received for publication May 10, 1966. Supported by Public Health Service Research Grant MH 05306 and MH 10094.

[1] For convenient reviews, see Kurt Koffka, *Principles of Gestalt Psychology,* 1935, 423-528; J. P. Guilford, *Psychometric Methods,* 1954, 302-340.

[2] See J. G. Beebe-Center, *The Psychology of Pleasantness and Unpleasantness,* 1932; R. S. Woodworth, *Experimental Psychology,* 1938, 368-391, 448-449.

[3] Beebe-Center, *op. cit.,* 196.

[4] T. H. Choo, A study of time error in esthetic judgment of auditory stimuli, *Studies in Psychology, Ewha University,* 1, 1954, 67-83.

[5] Gilbert Brighouse and S. D. Koh, The time error in esthetic judgment, *Amer. Psychologist,* 5, 1950, 317. (Abstract)

EXPERIMENT I

Problem. The experiment aims to examine the effect of preference-level and range of stimuli on appearance of the *TE*. Although most of the studies in psychophysics report negative *TE*, it is also known that, when the stimulus-range is sufficiently wide, the *TE* becomes positive with the stimuli of lower intensity and negative with the stimuli of higher intensity. In a comparison of weights paired with a varying standard, for instance, Woodrow found a *TE* of zero in the region of 130 to 140 gm., negative *TE* above this region, and positive *TE* in the region below.[7] This trend was verified by Needham for loudness and by Bartlett for weight.[8] In Experiment I, it is similarly hypothesized that the second of the paired musical excerpts will be more preferred (negative *TE*) if the excerpts are pleasant and will be less preferred (positive *TE*) if they are unpleasant, and that this preference bias will increase as the pleasantness or unpleasantness of the pair increases and will be negligible for the indifferent pairs. This analogy also is reasonable because affective judgments have repeatedly demonstrated the features of prothetic or intensity continua.[9]

Method and procedures: (1) Subjects. The *Ss* for the comparison of vocal selections consisted of a total of 159 college students, 92 hospital aides, and 118 men hospitalized for alcoholism; and the *Ss* for the comparisons of piano selections were 23 college students, 46 student nurses, 81 hospital aides, and 48 men, alcoholic patients. The college students were from Jamestown College and the University of North Dakota, while the student nurses, aides, and alcoholics were from the Jamestown State Hospital, North Dakota. The college students and the student nurses were mostly freshmen and sophomores, with an average age of 20 yr. The average age for the hospital aides was 44 yr. and that of the alcoholic patients was 42. The average education was 10 yr. for both of these groups. About 90% of the patients were semi-skilled workers, farmers, and laborers before their hospitalization. No psychopharmacological treatment was given to these patients, and no patient with psychotic or organic complications was included.

---

[6] For reviews, see Guilford, *op. cit.*, 305-311; Woodworth, *op. cit.* 438-449.

[7] Herbert Woodrow, Weight discrimination with a varying standard, this JOURNAL, 45, 1933, 391-416.

[8] J. G. Needham, Contrast effects in judgments of auditory intensities, *J. exp. Psychol.*, 18, 1935, 214-226; R. J. Bartlett, Measurement in psychology, *Advancement of Science*, 1939, 1-20.

[9] For a review, see Gösta Ekman and Lennart Sjöberg, Scaling, in P. R. Farnsworth (ed.), *Annual Review of Psychology*, 16, 1965, 451-474.

(2) Stimuli. A preliminary experiment designed to derive the affective scale values of the selections used has been reported elsewhere.[10] In brief, the stimulus-pool consisted of 51 vocal and 60 piano excerpts which were separately scaled on the basis of 9 ordered categories, ranging from 'most pleasant' (9), through 'indifferent' (5), to 'most unpleasant' (1). One set of 15 vocal pairs and another set of 20 piano pairs were made for the present study by selecting 30 vocal excerpts and 40 piano excerpts, respectively, from the pool and recording them on tape. These excerpts were approximately evenly spaced over the entire affective range of the stimulus-pool and had comparatively small variability. The two excerpts of each pair had approximately equal affective values, the pleasant being paired with the pleasant and the unpleasant with the unpleasant. The scale difference between pairs of excerpts was less than one point on the nine-point scale. The material was presented with intrapair and interpair time-intervals of about 6 and 10 sec., respectively. Each vocal piece lasted about 60 sec., and each piano piece about 15 sec. Each tape included 2 practice pairs.

(3) Procedures. The *S*s were instructed to compare the two paired excerpts on the basis of their impression of pleasantness-unpleasantness, checking one of the seven ratings appearing on the answer form. These ranged from "the second is extremely more pleasant than the first" (7), through "equal to the first" (4), to "extremely less pleasant than the first" (1). The *S*s were advised, however, to avoid the category of equal as much as possible. The 15 vocal and 20 piano pairs were each presented to the *S*s twice, once in AB order and then in BA order, by a tape-recorder. That is, each excerpt in a pair was presented once first and once second. A negative *TE* occurred whenever an excerpt of a pair was more preferred in the second than in the first presentation, since the preferencial bias tended toward overestimation of the second presentation or underestimation of the first presentation. The opposite tendency constituted a positive *TE.* To another group comparable to the above, the pairs were presented in BA order first and then in AB order. From the AB order of the first group and BA order of the second group, it was also possible to obtain the *TE.* This between-group *TE* should eliminate any effect of repetition that might exist in the within-group comparison obtained from the same group to whom the pairs were presented twice, in opposite orders.

In the treatment of the data, it was possible to compute two kinds of *TE*'s, the 'percentage *TE*' and the 'deviational *TE*.' In the former, the three categories of more pleasant and the three of less pleasant were combined into the categories *P* and *U*, respectively, while the equal ratings were divided proportionally into *P* and *U*. The *TE* was calculated by subtracting the percentage of an excerpt in the second presentation from percentage of the same excerpt in the first presentation. This *TE* is, therefore, close to the traditional use of D% for the *TE.*[11] The deviational *TE*, on the other hand, is defined as the number of categorical shifts from the comparative preference for the AB order to that for the BA order, or vise versa. That is, for example, if an excerpt is judged 'mildly more pleasant' in AB order, and 'extremely more pleasant' in BA order, a shift of two categories has occurred. This computational method should provide, at least, an ordinal scale for the *TE*. Although the results of the present study are primarily presented in terms of the percentage *TE*, some of that statistical analyses were performed with the deviational *TE*.

---

[10] S. D. Koh, Scaling musical preferences, *J. exp. Psychol.*, 70, 1965, 79-82.
[11] See Guilford, *op. cit.*, 306.

Results and discussion. Fig. 1 and 2 summarize the results from the vocal and piano comparisons, respectively, in terms of the percentage *TE*. Since the affective values of the paired excerpts were not the same among the groups, the *TE* from each group was plotted against the scale-values which were derived from the comparable group in the previous study.[11] These results are conclusive regarding the appearance of the affective *TE* in general and the effect of preference level and range on the effective *TE* in particular, confirming the hypothesis stated. The rank order correlations between the percentage *TE* and the affective scale values of the paired excerpts are shown in the upper right corner of each figure (figures with decimal points). All correlations show the negative trend of the *TE* as the scale-value for preference increases, although two of them (0.45 and 0.48 of Fig. 1) fail to meet a significance level of 0.05. The regression of the *TE* over the scale-value was computed by the least-squares method, and the regression-line and the regression-coefficient (lower figures) are also shown in each figure. Since the preference-value of 5 is 'Indifferent,' defined by the comparable group in the previous study, the regression-line should pass from positive to negative at this value. When the preferential value 5 is substituted in the regression-equation, the *TE* for each group ranges from −15% to 3% (average, −5.2%) in vocal comparisons and from −14% to 3% (average, −3.9%) in piano comparisons. When entire *TE* ranges of approximately ±60% to ±80% are taken into consideration, these outcomes appear to be close to zero and, accordingly, to the prediction.

The correlations and regression coefficients for the college groups are considerably smaller than those for the other groups, in both piano and vocal comparisons. The student's ability to identify the individual excerpts might be one of the reasons, since if an excerpt is familiar it will be selected or rejected no matter what is the order of presentation. This is especially possible with the vocal excerpts which are longer than the piano excerpts, and contain more information through verbal meaning and instrumental accompaniment. The slopes of the regression-lines in vocal comparisons are, in fact, consistently less steep than those in piano comparisons.

The *TE*'s from the alcoholic patients reveal essentially the same features as those from other *Ss*. When, however, compared to the hospital aides (men), who are similar in age, education, and socio-economic background, the alcoholics tend to

---

[11] Koh, *op. cit.*

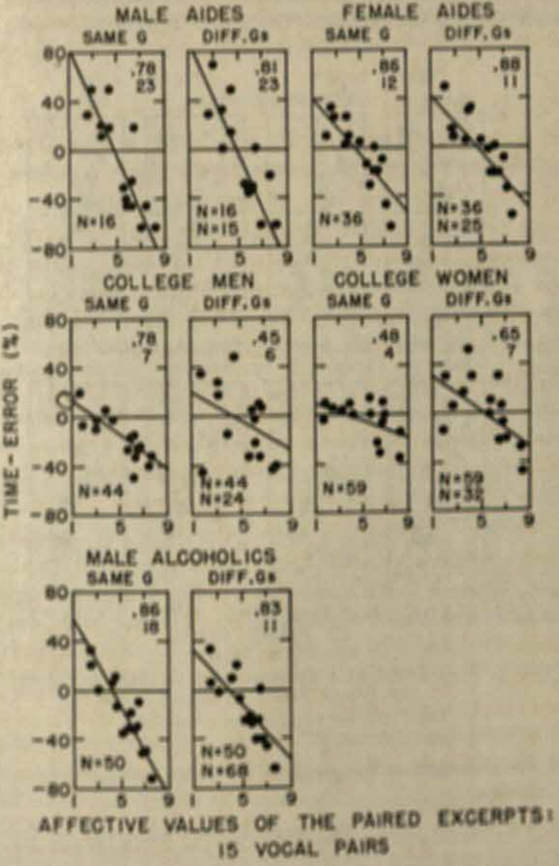


FIG. 1. TIME-ERROR AS A FUNCTION OF AFFECTIVE VALUE OF PAIRED VOCAL EXCERPTS

Same G and Diff. Gs indicate respectively, within-group *TE* and between-group *TE*. Figures in the upper right corner are rank order correlations and regression coefficients (with negative signs omitted). Affective values 1, 5, and 9 are respectively 'most unpleasant,' 'indifferent,' and 'most pleasant.'

show less positive *TE*'s (not significant) and smaller regression coefficients. The differences of the regression coefficients between these two groups in the between-group *TE* for the vocal excerpts ($t = 2.8$, $df = 26$) and that for the piano excerpts

(t = 2.17, df = 36) are significant beyond the 0.05 level. There is some evidence which correlates the *TE* to personality variables or pathological classifications.[13] Since the basic personality and pathological characteristics of alcoholics tend to be a mixture of various patterns, it would be premature to make a simple prediction concerning the *TE* preferred by the alcoholics. The smaller regression, *i.e.* the smaller *TE* tendency of the alcoholics as compared to the aides seems, however, worthy of further confirmation.

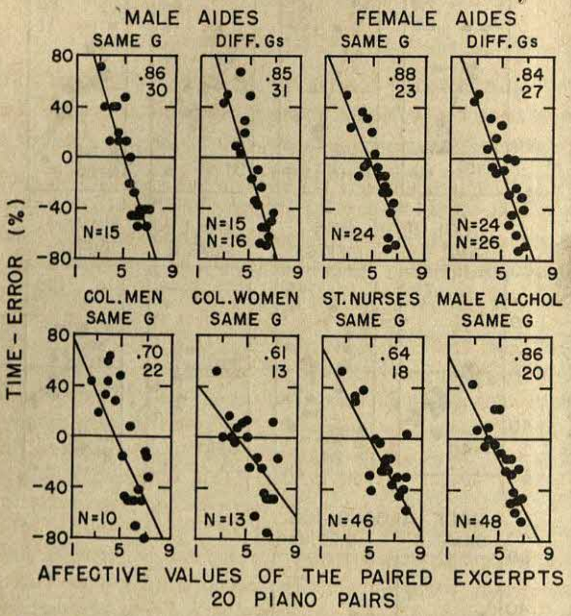


FIG. 2. TIME-ERROR AS A FUNCTION OF AFFECTIVE VALUE OF PAIRED PIANO EXCERPTS

See Fig. 1 for explanation of symbols.

The present results clarify the inconsistency observed in the previous studies on the affective *TE*. The predominantly negative *TE* reported by Choo in comparisons of melodies and by Brighouse and Koh in comparisons of paintings could be expected if the simple tones and rhythms in the former and colors and designs in the latter were on the whole less pleasant than the melodies and paintings, respectively.[14]

---

[13] On personality variables, see P. S. Holzman, The relation of assimilation tendencies in visual, auditory, and kinesthetic time-error to cognitive attitudes of leveling and sharpening, *J. Pers.*, 22, 1954, 375-394; P. S. Holzman and G. S. Klein, Cognitive system-principles of leveling and sharpening: Individual differences in assimilation effects in visual time-error, *J. Psychol.*, 37, 1954, 105-122; and Riley Gardner, P. S. Holzman, G. S. Klein, Harriet Linton, and D. P. Spence, Cognitive control, *Psychol. Issues*, 1, 1959, 22-30. On pathological classification, see Gordon Claridge, The excitation-inhibition balance in neurotics, in H. J. Eysenck (ed.), *Experiments in Personality*, 1960, 107-154; H. G. Birch, Ira Belmont, and Eric Karp, Excitation-inhibition balance in brain-damaged patients, *J. nerv. ment. Dis.*, 139, 1964, 537-544; and The prolongation of inhibition in brain-damaged patients, *Cortex*, 1, 1965, 397-409.

EXPERIMENT II

Problem. With the verification of the affective *TE* as a function of stimulus-level and range, Experiment II aims to assess the 'background' effect on preferential comparisons.

Guilford and Park interpolated a heavier or a lighter weight between a standard and its comparison weight, and found a shift of the point of subjective equality, resulting in a more positive *TE* with heavier interpolation and a more negative *TE* with lighter interpolation.[15] The experiments by Lauenstein both with vision and audition,[16] by Pratt with audition,[17] and by Holzman with vision, audition, and kinesthesis confirmed this trend.[18] These studies have special significance, since they were primarily conducted to test the theories advanced for the *TE* phenomena and comparative judgment in general. There has been no attempt, however, to assess the effect of interpolation on preference comparisons with an experimental paradigm comparable to the above. The experiments performed in connection with the law of effective contrast, however, have some relevance to the problem, although these studies dealt with absolute judgments rather than comparative judgments. Bacon, Rood, and Washburn interpolated a very agreeable or disagreeable colored paper between each two of medium colors and found that the pleasure of an agreeable experience is heightened or lowered by an immediately preceding disagreeable or agreeable experience, respectively.[19] Harris found a mass contrast effect from the whole series of interpolated colors rather than the individual contrast effect.[20] Beebe-Center confirmed this mass effect, using olfactory stimuli and distributing the experimental session over several days.[21]

In Experiment II, it is hypothesized that when piano excerpts of medium preference are paired for affective comparisons, the *TE* will tend to be negative if there is interpolated within each pair an excerpt of low preference and it will tend to be positive if there is interpolated an excerpt of high preference. Two experiments were conducted in the study.

---

[14] Choo, *op. cit.;* Brighouse and Koh, *op. cit.*

[15] J. P. Guilford and D. G. Park, The effect of interpolated weights upon comparative judgments, this JOURNAL, 43, 1931, 589-599.

[16] Otto Lauenstein, Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler, *Psychol. Forsch.*, 17, 1932, 130-177.

[17] C. C. Pratt, Time-error in psychophysical judgments, this JOURNAL, 45, 1933, 292-297.

[18] Holzman, *op. cit.*

[19] M. M. Bacon, E. A. Road, and M. F. Washburn, A study of affective contrast, this JOURNAL, 25, 1917, 190-291.

[20] A. J. Harris, An experiment on affective contrast, this JOURNAL, 41, 1929, 617-624.

[21] Beebe-Center, The law of affective equilibrium, this JOURNAL 41, 1929, 64; *op. cit.*, 1932, 222-237.

Method and procedures: (1) Subjects. The *S*s for the first experiment were 43 student nurses from the Jamestown State Hospital, and the *S*s for the second experiment were 34 men and 56 women from Jamestown college (Jamn) and the University of North Dakota (UND). They were mostly freshmen and sophomores.

(2) Stimuli. In the first experiment, 10 pairs of piano selections of medium preference (mean scale values $= 5.8$, $\sigma = 0.51$) were selected from the 20 pairs used in Experiment I. Five of these medium 10 pairs had interpolated excerpts of high preference (mean $= 7.6$, $\sigma = 0.03$), and in the other five pairs there were interpolated excerpts of low preference (mean $= 3.3$, $\sigma = 0.45$). These 10 pairs, including the interpolated excerpts, were recorded on tape in AIB order and then in the reversed BIA order. The recording order of the pairs with high and low interpolation was mixed. The three excerpts in each pair, AIB or BIA, were separated by about 6 sec., and the pairs were separated by about 10 sec.

After the first experiment, it was felt that the results would be more conclusive if in the same medium pairs there were interpolated high excerpts for one group and low excerpts from another group. In the second experiment, therefore, 10 pairs of medium preference for college students (mean $= 5.1$, $\sigma = 0.57$ for men; mean $= 5.0$, $\sigma = 0.85$ for women) were made from the stimulus-pool of 60 piano selections described in Experiment I. On the first tape in 5 of these 10 pairs there were interpolated excerpts of high preference (mean $= 7.1$, $\sigma = 0.18$ for men; mean 7.5, $\sigma = 0.13$ for women), while in the other five pairs there were interpolated excerpts of low preference (mean $= 3.0$, $\sigma = 0.39$ for men; mean $= 3.3$, $\sigma = 0.45$ for women). On the second tape, the high and low interpolations of the first tape were exchanged. That is, in the five pairs with high preference excerpts on the first tape there were interpolated low preference excerpts on the second tape, and vice versa. Other stimulus features of these two tapes were the same as those in the first experiment.

(3) Procedures. The instructions to the *S*s in both experiments were the same as in Experiment I, with the exception that the *S*s were asked to ignore the middle or interpolated excerpt in their comparisons of their preferences for the first and the third. The computation of the *TE* was the same as that from the same group in Experiment I.

Results and discussion. The left hand portion of Fig. 3 summarizes the results of the first part of Experiment II. The percentage of *TE*s from the high interpolation (filled circles) are, on the whole, found above those from the low interpolation (open circles). If the sums of the deviational *TE* of each *S* are computed for the five pairs with high interpolation and for the five with low interpolation, the average difference between these sums is 3.65 categories (0.73 categories for each pair) in the predicted direction, and this shift is significant beyond the 0.0001 level ($z = 5.21$). An inspection of Fig. 3, however, reveals a confounding of the two variables. There is an effect of stimulus-level and range that overlaps with the effect of the interpolation. The former, which gives a more negative *TE* as preference increases, is a confirmation of the finding in Experiment I.

As mentioned above, the second part of Experiment II was designed to control the effect of stimulus-level by interpolating the high and low excerpts in the same pairs for two comparable groups. Fig. 4 shows the results of this second experiment. The percentage *TE* are presented separately for men and women and also for each set of five pairs where the high and low interpolation were exchanged for the two independent groups (Jamn and UND). Each vertical line represents a pair of selections. Among the total of 20 pairs, all but one show the effect of interpolation as predicted; that is, the trend toward positive *TE* with high interpolation and toward negative *TE* with low interpolation. In terms of the deviation *TE*, the average difference between the five high and five low interpolations range from 4.0 to 5.1 categories, and these differences are significant beyond the 0.025 level (men, $df = 32$, $t = 5.34$, 2.12; women, $df = 54$, $t = 3.79$, 2.86; one-tailed). If the deviational *TE*'s for

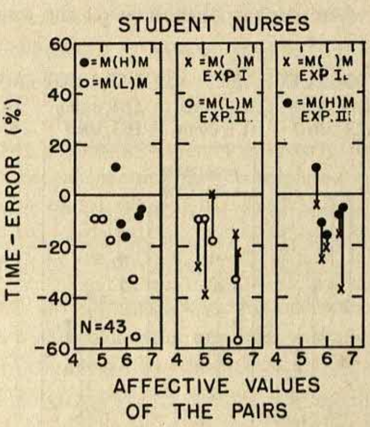


FIG. 3. TIME-ERROR WITH INTERPOLATION OF OTHER EXCERPTS WITHIN PAIR BEING COMPARED
'M(H)M' and 'M(L)M' indicate medium pairs with pleasant and unpleasant interpolation, respectively. Left figure shows results from Experiment II, and the middle and right figures show comparisons between Experiment I and Experiment II. 'M( )M' indicates no interpolation.

the five high and five low interpolation within the same group (Jamn or UND) are computed as in the first experiment, the *TE* shift due to the interpolation ranges from 2.5 to 5.5 categories and these shifts are again significant beyond the 0.025 level (men, $df = 16$, $t = 3.48$, 2.26; women, $df = 27$, $t = 4.06$, 4.22; one-tailed).

Fig. 4 again demonstrates the negative trend of the *TE* as the preference increases, in spite of the narrow affective range. If the regression-lines for the high (filled circles) and low interpolations (open circles) are separately drawn by the graphic method, excepting the right figure (women: *H*, UND; *L*, Jamn), the distance between the two points where the two lines intersect the zero *TE* is roughly one category step. This means that the point of subjective equality is shifted in the predicted direction roughly a half category step by the high and low interpolation, respectively. The extremely narrow preference range might obscure this trend in the right hand portion of Fig. 4.

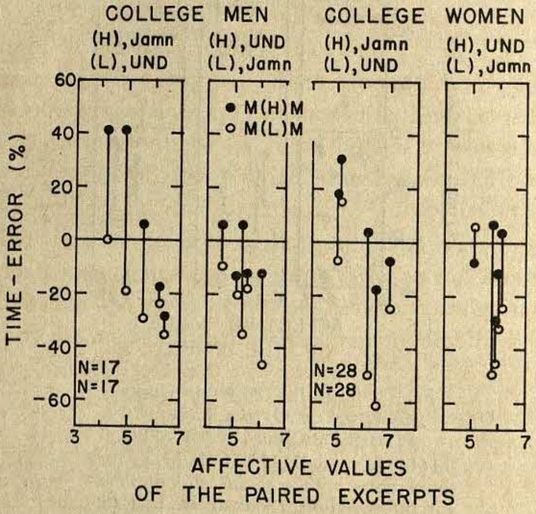


FIG. 4. TIME-ERROR WITH INTERPOLATION

See Fig. 3 for explanation of symbols.

Another point of interest is to compare the results from the student nurses in Experiment I with those from the student nurses in the first part of Experiment II, both of which are comparisons of piano excerpts. As stated above, the 10 medium piano pairs of the present experiment were identical with 10 of the 20 piano pairs of Experiment I. In Experiment II, however, 5 of the 10 pairs had interpolated low affective excerpts and the other 5 had high excerpts, while there were no interpolations in Experiment I. It could be predicted that the *TE* from the 5 pairs with high interpolation of Experiment II would be more positive than the *TE* from the 5 corresponding pairs with no interpolation. The right-hand one-third of Fig. 3 shows that the results were all in the predicted direction, although the difference in terms of the deviational *TE* for the 5 pairs combined is not significant ($df = 87$, $t = 1.64$). On the other hand, the comparisons between the low interpolation of Experiment II and the no interpolation of Experiment I for the 5 other pairs are largely inconsistent. This is shown in the middle panel of Fig. 3. Only three of the five pairs of the present comparisons show the expected trend.[22] Since Experiment I included 10 more pairs than the 10 pairs just discussed, and the intrapair interval is different between Experiments I and II, the above comparison is not too conclusive.

EXPERIMENT III

Problem. The background or interpolation effect was examined further through another experimental design. If the preferred *TE* is a function of the affective value of the interpolated material, it can be further hypothesized that the *TE*, both positive and negative, will be larger when the difference between the paired and the interpolated excerpts is large than when the difference is small. That is, if pairs both of which have a high affective value have inserted an excerpt with low affective value, while pairs of low affective value are combined with high affective excerpts, the regression of the *TE* over the affective level of the pairs will be steeper than if high and low affective pairs are combined, respectively, with high and low affective insertions. In the former instances, the effects of preference level and interpolation would work in the same direction and maximize the appearance of *TE,* whereas in the later instance, only the effect of preference level would operate.

Method and Procedures: (1) Subjects. The subjects were 72 student nurses from the Jamestown State Hospital.

(2) Stimuli. From the 60-stimulus pool described in Experiment I, two tapes were made, with maximal and minimal differences. On the tape with the maximal difference, H(L)H and L(H)L, 10 pairs were recorded in the following ways. Four high affective pairs (scale-values 7.4 to 7.6) had low affective excerpts (3.5 to 5.7) interpolated, four low affective pairs (3.7 to 4.2) had high affective excerpts (6.4 to 7.6) interpolated, and two pairs of medium preference were combined with medium affectively interpolations, which would be located within the transitional region between the positive and negative *TE*s. The difference affective value between the eight paired excerpts and their interpolated excerpts, excluding the medium pairs, was 2.90 on the average. These 10 pairs were recorded twice on tape in a random order, first in AIB order and then BIA order. On the tape with minimal differences, H(H)H and L(L)L, three high affective pairs (all, 7.4), four medium affective pairs (5.5 to 6.5), and three low affective pairs (3.1 to 4.2) were prepared with interpolations of approximately the same affective value. The scale-difference between the 10

---

[22] Pratt, in loudness comparisons, found larger negative TE with low interpolation than with no interpolation. Pratt, *op. cit.,* 292-297.

paired excerpts and their interpolated excerpts was 0.85 on the average. Other features of the first and second tapes were the same as with the tapes of Experiment II.

(3) Procedures. The first tape was played to one group (N = 39) and the second tape was played to another group (N = 33). The instruction to the *Ss* and the computation of the TE were the same as those in Experiment II.

Results. The results are summarized in Fig. 5. The trend toward negative *TE* as the affective values of the pairs increase or toward positive *TE* as the values decrease is again demonstrated in both of the groups. To compare the slope of the regression of the *TE* over the affective level of the pairs, the line of best fit was computed by the method of least squares separately for maximum (open circles) and minimum (filled circles) interpolations. The *TE* is related to the affective value of the pair *A* by the following equations: for maximal difference, $TE = 132.6—25.3\ A$; for minimal difference, $TE = 73.0—14.8\ A$. The regression line for the larger differences is steeper than that for the smaller differences, as predicted, and the difference between the two regression coefficients is significant beyond the 0.025 level ($df = 16$, $t = 2.40$, one-tailed). The scatter of the *TE* in Fig.

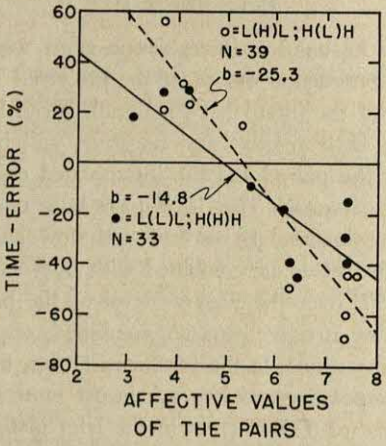


FIG. 5. TIME-ERROR WITH INTERPOLATION

'L(H)L' and 'H(L)H' indicate, respectively, the unpleasant pairs with pleasant excerpt interpolated, and the pleasant pairs with unpleasant excerpt interpolated. 'L(L)L' and 'H(H)H' indicate unpleasant and pleasant excerpts interpolated unpleasant and pleasant pairs, respectively.

5 shows a clear differentiation in the region of high affective value, H(L)H vs. H(H)H, but is rather ambiguous in the lower region, L(H)L vs. L(L)L.

Discussion and Conclusion

The behavior of affective *TE,* both as a function of level of preference and range (Experiment I) and of the 'background,' or interpolation, (Experiment II and III) in successive comparisons of pairs of musical excerpts is on the whole reliably demonstrated in the present study. It behaves much in the manner observed in psychophysics. The present findings, therefore, imply that affective judgment shares common underlying processes with psychophysical judgments, and might be explained by the related theories. Among the theories advanced, the theory of assimilation and of adaptation-level seem to account for the present results most successfully.[23] According to adaptation-level (*AL*), the basis of judgment is provided by the *AL,* which is a weighted average of all the past stimuli. If the *AL* theory is applied to the data of Experiment I, the second excerpt is compared not just with the first, but with the *AL* toward which the first excerpt is being pulled during the intrapair interval. That is, if the preference of the second excerpt is below the *AL,* it will tend to be judged less pleasant and if its preference is above the *AL,* it will tend to be judged more pleasant. The *TE,* accordingly, will be positive in the former (unpleasant pair) and negative in the latter comparisons (pleasant pair), and the trend will become more evident as the difference in preference between the paired excerpts and the *AL* increases. The rank order correlations and regression coefficients presented in Experiment I agree with this explanation. Furthermore, the prediction of zero *TE* at the scale-value 5, where the discrepancy from the *AL* is supposed to be small, was also generally confirmed in Experiment I. The *AL* theory is also consistent with the effects of background of interpolation found in Experiments II and III, since, as the theory postulates, the high or low interpolation would pull the *AL* upward or downward, respectively during the intrapair interval. Although the comparison between 'low' interpolation of Experiment II and 'no' interpolation of Experiment I was only a by-product of the present study, a separate experiment is needed to clarify the difference, in view of Pratt's critical experiment on the problem.[24] The issue seems to be whether the no interpolation corresponds to the absolute threshold as Pratt appears to believe, or corresponds to the *AL.* If the latter is the case, Pratt's observation of large negative *TE* with low interpolation rather than no interpolation is explainable from the *AL* theory.

An explanation similar to the *AL* theory could be formulated from the assimilation-theory, which would postulate the assimilation of the electro-chemical trace formed by the stimuli to the entire trace system in the brain as the basis of affective comparison. The *AL* theory, however, seems to incorporate this Gestalt thinking without committing itself to the electro-chemical processes assumed in the brain and to make the variables involved, especially the entire trace system, more explicit and concrete than the assimilation theory. Recently, Claridge has applied Eysenck's theory of cortical excitation and inhibition to the psychophysical *TE*.[25] The major difficulty with this theory as it stands is the fact that it does not satisfactorily explain the appearance of positive *TE*. To explain positive *TE* the theory has to provide some basis for an enhancement of the first excitation during the intrapair interval or else for a suppression of the second excitation. Birch, Belmont, and Kapp observed a tendency toward positive *TE* in auditory comparisons by the brain-damaged patients and attributed the cause to the prolongation of inhibition or of recovery from the first excitation.[26] The cortical excitation-inhibition theory may gain some support from such observations, but the reason that the positive and negative *TE* resulted with unpleasant and pleasant excerpts, respectively, within the same person is still difficult to derive from the theory.

---

[23] See Lauenstein, *op. cit.,* and Koffka, *op. cit.,* 423-481, for the theory of assimilation of traces. Also, Helson, *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior,* 1964.

[24] Pratt, *op. cit.*

The finding that the dimensional nature of affective judgment is prothetic rather than metathetic is further confirmed in the present study. Among Stevens' four empirical criteria which distinguish the two kinds of continua, the nonlinear or logarithmic relation between the category and magnitude scales has been repeatedly demonstrated in affective judgments.[27] The appearance of the *TE* is another criterion for prothetic continua and is consistently shown in the present study.

The observation that the size of the *TE*, both positive and negative, tends to increase proportionally as the affective values of the musical pairs deviate from the indifferent region is interesting. Since the *AL* theory takes the indifferent point as a functional zero in psychological continuum, this trend seems to follow Weber's generalized law. Luce and Galanter have suggested that the constant error like the *TE* may follow Weber's law when the *jnd* does.[28] Furthermore, Ekman and his associates have demonstrated that on prothetic continua the subjective variability tends to grow as a linear function of the subjective magnitude.[29] This discriminative imbalance over the continuum, which is reflected by the j.n.d., constant error, and subjective variability, may be the major cause of the *TE* phenomena, as Stevens speculates.[30]

SUMMARY

The behavior of the affective time-error (*TE*) as a function of stimulus-level and range and as a function of an interpolated stimulus in comparisons of preferences for short musical excerpts was examined, using various groups of college students, hospital aides, and alcoholic patients as *S*s. The second excerpt of each pair was increasingly less (positive *TE*) or increasingly more (negative *TE*) preferred, respectively, as the preference-level

---

[25] Claridge, *op. cit.*

[26] Birch, Belmont, and Karp, *op. cit.*, 1964 and 1965, 537-544 and 406-408.

[27] S. S. Stevens, On the psychophysical law, *Psychol. Rev.*, 64, 1957, 153-181. As applied to affective judgment, see Ekman and Sjöberg, *op. cit.*, 464; and Koh, *op. cit.*, 79-82.

[28] R. D. Luce, and Eugene Galanter, Discrimination, in R. D. Luce, R. R. Bush, and Eugene Galanter (ed.), *Handbook of Mathematical Psychology*, 1, 1963, 191-243.

[29] For a summary see Stevens, A metric for the social consensus, *Science*, 151, 1966, 523-534.

[30] Stevens, *op. cit.*, 1957, 156-159.

of the paired excerpts decreased, or increased, from the indifferent region. In the mid-region the *TE* became negligible. The rank order correlations between the preference level and size of the *TE* ranged from —0.45 to —0.88 (Experiment I). When in the indifferent pairs were interpolated pleasant excerpts or unpleasant excerpts, the *TE* tended to be positive in the former and negative in the latter case (Experiment II). When the difference in preference between the paired excerpts and interpolated excerpt was greatest the regression of the *TE* over preference level became significantly steeper than when the difference was least (Experiment III). The striking similarity of these observations to the psychophysical *TE* suggests common underlying processes between affective and psychophysical judgments and, accordingly, the applicability of the theories advanced in psychophysics, especially adaptation-level theory, to affective judgments. The present study also reinforces the previous observations of the prothetic or intensive nature of affective judgments.

SPATIAL GENERALIZATION AS A FUNCTION OF VARIATION IN A NONSPATIAL ATTRIBUTE, INSTRUCTIONS, AND INDIVIDUAL STRATEGIES

By H. WAYNE LUDVIGSON, Texas Christian University,
and WILLIAM F. CAUL, University of Pittsburgh

The following experiments investigated the spatial generalization of a voluntary response in a multiple probability-learning task presented to *S* as a series of 'horse-races.'[1] Compared with the more frequently studied kinds of generalization, this type of spatial generalization appears to involve relatively complex information-processing and hypothesis-testing behavior. Seven lamps, representing horses, are arrayed on a horizontal line. The lighting of a lamp is a cue for *S* to predict whether that particular horse will win or lose the current race. The basic finding is that if the symbolized horse in the middle of the array wins a large proportion of its races, 80%, while the lamps to either side all win a small proportion of their respective races, 20%, then the frequency with which *S*s predict 'win' to the side lamps is an increasing function of the nearness of the lamp to the middle.

GENERAL METHOD

Apparatus. The apparatus was similar to that used by Brown, Clarke, and Stein.[2] It consisted of a curved black panel on which seven lamps (7.5 w.) were mounted equally spaced in a horizontal row at 8° intervals. The lamps were approximately at eye-level 5 ft. from *S*'s face. For convenience, the middle lamp will be designated 0. The lamps to the right of 0 will be called +1, +2, and +3 and the lamps to the left will be designated −1, −2, and −3. Directly below the middle lamp were two 'win' and two 'lose' signs. One pair indicated *S*'s response, made by pressing a button mounted on the arm of his chair, and the other pair indicated the outcome of the trial. The room was illuminated by a small light behind *S*.

Procedure. The *S*s were told that "this is an experiment to determine how people make predictions in a simulated horse-race betting situation." The seven lamps were then explained as representing seven different race horses, each of which races against other imaginary horses that are not represented by lamps on the panel. The instructions emphasized that the outcomes of the races for one horse were completely independent of the outcomes for another horse.

---

* Revised manuscript received June 29, 1966. This research was supported by a grant from the Carnegie Corporation while the authors were at Carnegie Institute of Technology.

[1] J. S. Brown, F. R. Clarke, and Larry Stein, A new technique for studying spatial generalization with voluntary responses, *J. exp. Psychol.*, 55, 1958, 359-362.

[2] Brown, Clarke, and Stein, *op. cit.*, 361.

When a lamp came on, *S*'s task was to guess whether the horse would win or lose. Two seconds after he responded, *S* was informed of the outcome. Three seconds later all the lamps were turned off. *S*'s guess and latency of response were then recorded and the next trial begun. The latency measure was the elapsed time between the lighting of a lamp and *S*'s response.

During an experimental session, *S* was given 210 trials consisting of six blocks of 35 each. Within each block of 35 trials, each lamp was presented five times in randomized order with the restriction that a lamp could appear only once in three successive trials. A different order was used for each block and the sequence of blocks was varied from *S* to *S*. Within each block the prearranged percentage of reported wins and losses was maintained.

EXPERIMENT I

The first study sought to determine whether the amount of generalization depends, as in certain simpler tasks, on the number of common characteristics of the stimuli.[3] On the basis of the available data and theory, it would be expected that if spatial generalization in this task results from a response-tendency's becoming associated with all the aspects of the middle stimulus, then generalization to adjacent stimuli should be reduced when these stimuli differ from the middle in color in addition to spatial position.[4]

Method. The *S*s, (16 men and 16 women) from classes in introductory psychology, were randomly assigned to each of two groups. For Group S, all lamps were the same color. For Group D, the center and end lamps were one color, while the remaining lamps were a different color. Color was manipulated by using white and red bulbs counterbalanced with respect to sex and lamp position. In both groups, the horse symbolized by the middle lamp won 80% of its races, while all side lamps won 20%.

Results. Employing the measure used by Brown, Clarke, and Stein,[5] the percentages of win responses elicited by the lamps over the last five blocks of trials are presented in Fig. 1. An analysis of variance was performed on the frequencies of win responses given to the lamps, excluding the middle lamp because of its different win percentage. The factors in the analysis were Groups S and D, (Lamp Pairs ±1, ±2, ±3), and Side of Array (left and right of the middle). Although Side of Array was significant, it did not appear important for the question under investigation, especially since all interactions with it yielded *F*s less than unity. This factor will not be discussed further.

---

[3] J. B. Fink and R. M. Patton, Decrement of a learned drinking response accompanying changes in several stimulus characteristics, *J. comp. physiol. Psychol.*, 46, 1953, 23-27; S. H. White, Generalization of an instrumental response with variation in two attributes of the CS, *J. exp. Psychol.*, 56, 1958, 339-343.

[4] C. L. Hull, *Principles of Behavior*, 1943, 209-215.

[5] Brown, Clarke, and Stein, *op. cit.*, 360.

The remainder of the analysis of variance supported the conclusion that, in spite of the attempt to replicate the procedures of Brown, Clarke, and Stein, Group S displayed no spatial generalization. In contrast, Group D, for which the middle and end lamps were *different* in color from the rest, showed symmetrical, decreasing gradients. According to the Scheffé test for *a posteriori* comparisons,[6] Group D made reliably more win

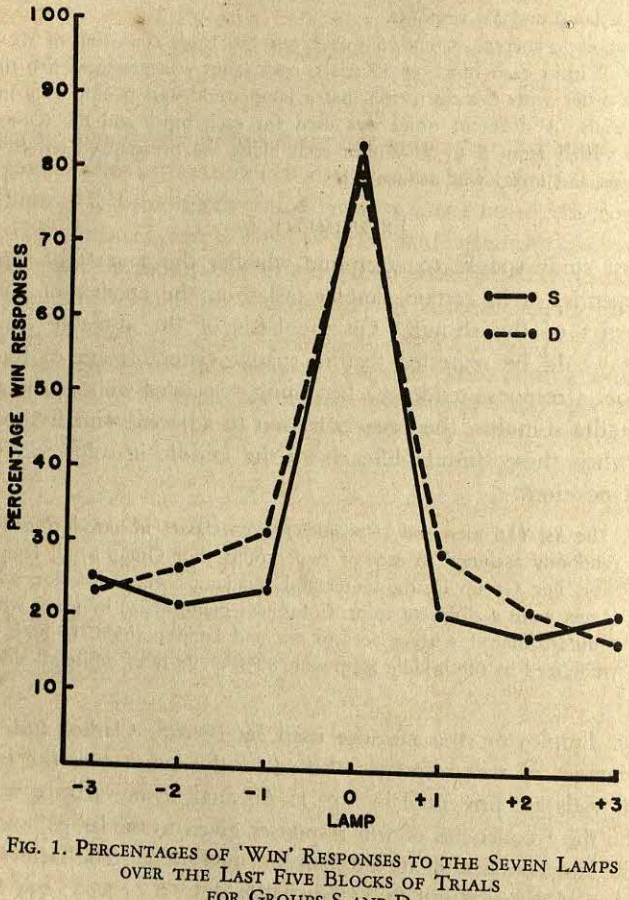


FIG. 1. PERCENTAGES OF 'WIN' RESPONSES TO THE SEVEN LAMPS OVER THE LAST FIVE BLOCKS OF TRIALS FOR GROUPS S AND D

responses to Lamps ±1 than to Lamps ±2 or ±3 ($p < 0.05$). Thus, the relationship between Groups S and D is exactly opposite to what would be expected were the amount of generalization an inverse function of the number of differing stimulus-attributes. This would appear to cast doubt on the proposition that spatial generalization, when it occurs, results from a response-tendency's becoming associated with all aspects of the middle stimulus and generalizing on the basis of all aspects.

---

[6] B. J. Winer, *Statistical Principles in Experimental Design,* 1962, 88.

EXPERIMENT II

The finding that decreasing the similarity of the stimuli actually increased spatial generalization indicated that further information regarding the role of the distinctively colored bulbs was needed. The bulbs may have served to direct attention toward certain lamps. Experiment II sought to determine whether similar effects can be accomplished through instructions which direct attention toward the end and middle lamps.

Method. Two groups, each comprised of eight men and eight women, received the same treatments as in Experiment I except for additional instructions. After the *S*s were instructed in the usual manner as to the independence of the horses and asked to try to remember how often each horse wins on its own track, they were then told:

> Actually this is kind of hard to do since there are so many horses. So what I want you to do is this: At first, try to get a pretty good idea of how often the two end horses and the middle horse win, and then later on after you have sort of figured out these horses, pay closer attention to the remaining horses. Of course, you'll have to pay some attention to all horses even at the beginning, but pay particular attention to the horses on each end and the middle one at first.

These 'attentional instructions' were given to a group for which all the lamps were the same color (Group S-AI) and to a second group (Group D-AI) which had distinctively colored lamps in the center and end positions in the manner of Group D of Experiment I. That is, all bulbs were white except for the middle and end bulbs of Group D-AI, which were red. Group S-AI was intended to ascertain whether instructions and distinctive lamps are functionally equivalent, while D-AI was included to determine whether the combination of distinctive bulbs and instructions would further augment the effect.

Results. An analysis of variance of the number of win responses over the last five blocks of trials indicated that Groups S-AI and D-AI did not differ significantly from each other and that no other factors interacted with Groups (all *F*s < 1). In view of this, the data for Groups S-AI and D-AI were pooled into a single group, designated AI.

To ascertain whether the instructions and distinctive stimuli function equivalently in the present task, Group AI was compared with Group D of Experiment I. Fig. 2 presents the data for Group AI and, for convenience of comparison, Group D. Analysis of variance similar to the previous one revealed a significant Groups X Lamp Pairs interaction ($p < 0.001$), indicating that instructions and distinctive lamps are not equivalent operations. This may be seen in Fig. 2. Group AI displays an upturning of the end points to which they were specially instructed to attend. No such upturning occurred in Group D which had distinctive bulbs in these end positions. In addition to the upturned ends, Group AI displayed gradients between Lamps ±1 and ±2, which is reliable using the Scheffé test ($p < 0.01$).

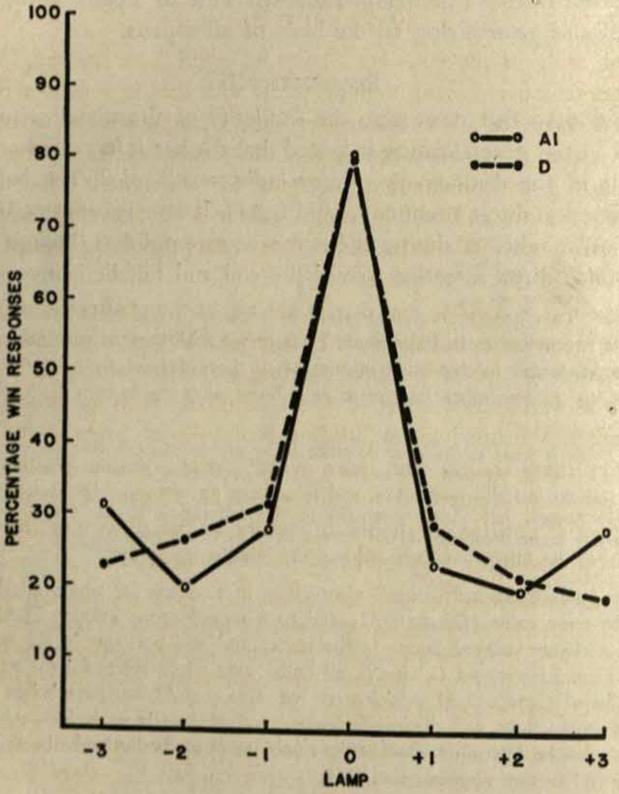


FIG. 2. PERCENTAGES OF 'WIN' RESPONSES OVER THE LAST FIVE BLOCKS OF TRIALS FOR GROUP AI, AND GROUP D OF EXPERIMENT I

It may be concluded, therefore, that both attention-directing instructions and distinctive lamps produced generalization where the standard procedures failed, but that the operations were not equivalent because the instructions increased the level of response to the end lamps, whereas the distinctive color by itself did not. Also, when instructions and distinctive stimuli were jointly imposed, the instructions were clearly prepotent, since the response was like the response to instructions alone.

EXPERIMENT III

The upturning of the ends for Group AI could have resulted from a kind of response-bias having nothing to do with the 80% winner in the middle. For example, given the instructions, *S* might attempt to predict the exact sequence of events for singled-out lamps, while, for the most part, predicting the predominant outcome for the others. This would tend to hold the end and middle points closer to a 50% level as compared with the other points. Were this the case, the effect should be present even in the absence of the 80% winner.

In addition, it is possible that part, if not all, of the gradient in response percentage around the middle might be attributable, not to generalization, but to an effect of 'spatial position' yielding a function similar at least in shape to the well-known effect of serial position.[7] Such an effect could easily result if *S* learned the win–lose probabilities of lamps at the ends of the array more readily than those in the center. Previous studies with this task have not controlled for this possibility. Again, if there is an effect of spatial position, it should be present in the absence of the 80% winner. This possibility was examined in Experiment III.

Method. Equal numbers of men and women were randomly assigned to three groups of 16 *S*s each. With the exception that all lamps won the same percentage of races, viz. 20%, these groups received treatments identical to Groups S, D, and AI of the previous experiments. All lamps were white except for the end and middle bulbs for the D condition, which were red.

Results. Using the same measure as in the previous experiments, there appeared to be no systematic trends across lamps, the values centering around a percentage of 24.5. An analysis with the three groups and the seven lamps comprising the levels of two factors revealed no evidence for differences among the Groups or Lamps, nor an interaction between these factors (Fs < 1). More detailed analyses substantiated the conclusion that there was no evidence of a response-bias or a spatial position effect. Thus the gradients and upturned ends of the previous experiment appear to depend upon the presence of a big winner in the middle.

---

[7] S. M. Ebenholtz, Position mediated transfer between serial learning and a spatial discrimination task, *J. exp. Psychol.*, 65, 1963, 603-608.

General Results and Discussion

The results of the three experiments may be summarized as follows: Variation in a nonspatial attribute and instructions that directed attention produced 'spatial' generalization, whereas the standard condition did not. In fact, the instructions fostered a second generalization-like phenomenon but reversed in direction. Moreover, there was no evidence of any predisposing bias toward a particular spatial position.

Although justified, these statements, if unqualified, obscure an important source of variance. It has been argued that in a decision-making task such as this, the *Ss* may operate under quite different strategies.[8] Some seem content simply to choose (predict) the more frequent event, a tactic that will keep the frequency of correct responses high. Others attempt to solve the problem of predicting the exact sequence. This latter strategy tends to keep the percentage of responses on the favored alternative less extreme than the former strategy and, since the sequence is

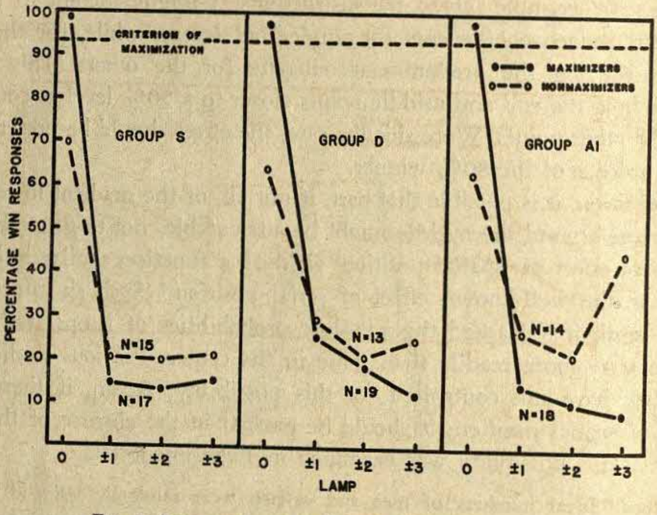


Fig. 3. Percentages of 'Win' Responses over the Last Three Blocks of Trials of *Ss* from Groups S, D, and AI of Experiment I and II Classified as Maximizers and Nonmaximizers

The number of *Ss* falling in each category is given on the figure. The criterion of maximization used to categorize *Ss* is also shown. Data for the lamps equidistant from the middle have been pooled.

---

[8] J. J. Goodnow, Determinants of choice-distributions in two-choice situations, this Journal, 68, 1955, 106-116.

unpredictable, to result in less than the maximal number of correct predictions.

It was decided to determine whether generalization in individuals who apparently employed a 'maximizing' strategy differed from individuals who employed other strategies. Persons in Experiments I and II who gave at least 14 'win' responses to the middle lamp out of 15 trials in the last three blocks were classified as 'Maximizers.' Any number less than this categorized them as 'Nonmaximizers.' Fig. 3 shows the percentages of win responses for these two categories of *S*s as well as the number of *S*s falling in each category. The three panels represent the three treatments of Experiments I and II. Response-latencies, heretofore not presented for the sake of brevity, are shown in Fig. 4 for these same *S*s classified in the same way.

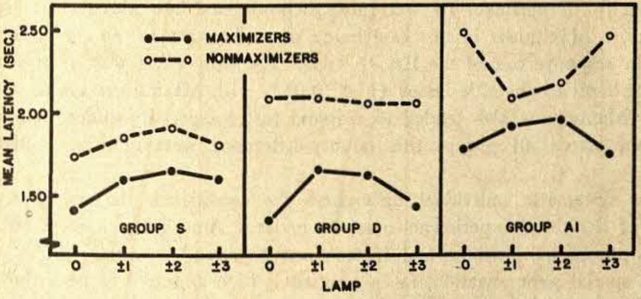


FIG. 4. MEAN RESPONSE-LATENCIES OVER THE LAST THREE BLOCKS OF TRIALS FOR MAXIMIZERS AND NONMAXIMIZERS OF GROUPS S, D, AND AI

There are three noteworthy features of these data. First, those individuals who maximized with respect to the middle lamp also tended to maximize with respect to the other lamps, giving mostly 'lose' responses. (The 'lose' outcome is, of course, the more frequent for these positions.) Conversely, Nonmaximizers consistently responded with less extreme percentages. Pooling across the three groups, the Nonmaximizers gave significantly more 'win' responses to the side lamps than did the Maximizers ($p < 0.001$). Gross variability among *S*s, then, is not attributable to differences in over-all relative preference for win or lose as a response, but rather to differences in the frequency with which the more likely event is predicted. As may be seen from Fig. 4, these differences are accompanied by consistent differences in response-latency. Maximizers were reliably faster responders to the side lamps ($p < 0.05$), as well as to the middle lamp ($p < 0.001$).

Secondly, in Groups D and AI, the trends in the percentage of win responses to the side lamps differ for Maximizers and Nonmaximizers. Considering the groups singly this interaction was statistically significant in Group AI ($p < 0.005$) but not in Group D. Similar analyses were performed on the trends of response latency for Groups AI and D. Although significant only at the 0.10 level, these tests, based on responses to the side lamps, do not take into account Lamp 0 which contributes markedly to the differences in trend.

Thirdly, whereas the percentages of win response of Groups D and AI appear to display two types of generalization, the latencies seem influenced by the processes underlying only one type. Spatial generalization does not seem to affect latency as evidenced (1) by the absence of a consistent difference in response to Lamps ±1 and ±2 between Groups D and AI, which display spatial generalization, and Group S, which displays no generalization, and (2) by the similarity of the latency function for Maximizers, who exhibit a simple spatial gradient, to the latency-functions of Experiment III in which there was no possibility of spatial generalization. Therefore, it seems likely that the processes underlying the second type of generalization in Groups D and AI not only produced more win responses to Lamps ±3 among Nonmaximizers but also slowed their responses to Lamp 0 and ±3.

The data for Experiment III were also dichotomized into Maximizers and Nonmaximizers, a Maximizer in this case being one who gave to the center lamp zero or one win response out of the last 15 trials. The only effect was in over-all level of win responses to the side lamps ($p < 0.01$), with Maximizers giving the fewer responses. Maximizers also tended to respond faster especially under standard conditions, but across all groups this latency-difference was significant only at the 0.10 level.

If these systematic individual differences are considered, the previously stated summary of the three experiments must be revised. Among Maximizers both variation in a nonspatial attribute and instructions to attend to certain lights resulted in simple spatial generalization, *i.e.* a gradient which appeared to be a linear function of spatial position. These same conditions had a more complex effect on Nonmaximizers, producing boh a spatial gradient from Lamps ±1 to ±2 and a second generalization-like phenomenon at Lamps ±3. Conditions S, D, and AI appeared to fall on a continuum with respect to the extent to which they produced this second, 'attention-mediated' generalization.[9] Viewed in this manner, it is tempting to conclude that the latter two experimental operations differ only in potency and not in kind—a conclusion considerably different from that suggested before *S*s were divided into Maximizers and Nonmaximizers.

A basic problem is why there was no evidence of spatial generalization in Group S, since the procedures were highly similar to those previously shown to give pronounced gradients. It is possible that the present instructions emphasized to a greater extent, or the *S*s accepted more readily, the independence of the lamp outcomes, with the result that spatial generalization was inhibited under the standard procedures. The distinctive bulbs and instructions for attention, in singling out more than one lamp, might have partially counteracted this set for independence. Should this be true, it would suggest that this spatial generalization, like that of a different voluntary response, depends upon a delicate balance of competing processes.[10]

---

[9] This second type of generalization might result from a mediational response or categorization which effects a similarity among the end and center lamps. It might be argued that the instructions and distinctive lamps also produced categorizing activity among the Maximizers, but of such a nature that the singled-out lamps acquired distinctiveness rather than similarity. Although the data do not require this latter assumption, the Maximizers' differential responsiveness to the end and center lamps, and consequently the spatial gradient, might have been enhanced by such a mechanism.

---

[10] J. S. Brown, E. A. Bilodeau, and M. R. Baron, Bidirectional gradients in the strength of a generalized voluntary response to stimuli on a visual-spatial dimension, *J. exp. Psychol.,* 41, 1951, 52-61.

JUDGMENTS OF GRAMMATICALNESS OF SOME ENGLISH SENTENCES AND SEMI-SENTENCES

By Lawrence E. Marks, New Haven, Connecticut

Strings of words that adhere completely to the syntactic and semantic rules of a language are grammatical sentences of that language. A simple distinction can be made between these grammatical sentences and all other, ungrammatical strings of words. Within the realm of ungrammatical strings, however, it is obvious that some have more structure than others (to use Chomsky's examples: "golf plays John" and "a the ago").[1] Many linguists recognize the existence of semi-sentences that lie between fully grammatical and completely ungrammatical strings, and these linguists believe that semi-sentences can be classified according to their degree of grammaticalness.[2]

There appear to be at least two distinct ways to assign levels of grammaticalness to deviant sentences. The first, which is essentially that taken by Chomsky, assigns a grammatical level to a semi-sentence according to the extent that the grammar of non-deviant sentences pertains to that sentence.[3] A second approach is to obtain judgments of grammaticalness from fluent native speakers of the language and to use these ratings in order to generate scales of grammaticalness. The position taken here is that the former approach is incorrect to the extent that the scales generated using that approach yield results different from those obtained using subjective judgments. After all, learning a language is a psychological process. The ways in which verbal materials are remembered and understood reflect psychological operations. The basic hypothesis is that levels of grammaticalness based upon subjective judgments are highly dependent upon psychological operations that occur when verbal material is processed, operations that may not be reflections of what have been described as purely formal aspects of the language.

Specifically, it is proposed that the degree of ungrammaticalness produced in a fully grammatical sentence by a particular type of distortion will depend not merely upon the extent that the grammar of non-deviant sentences pertains to that semi-sentence, but additionally upon the serial position in the sentence at which the

---

* Received for publication August 11, 1966. From the John B. Pierce Foundation Laboratory. A portion of this paper is based upon part of a Doctoral thesis submitted to the Department of Psychology, Harvard University. The research was completed at Harvard University, Center for Cognitive Studies.

[1] N. Chomsky, Degrees of grammaticalness, in J. A. Fodor and J. J. Katz (eds.), *The Structure of Language: Readings in the Philosophy of Language*, 1964, 386.

[2] *Ibid.*, 384-389; J. J. Katz, Semi-sentences, in Fodor and Katz, *op. cit.*, 400-416; P. Ziff, On understanding "understanding utterances," in Fodor and Katz, *op. cit.*, 390-399.

[3] Chomsky, *op. cit.*, 384-389.

distortion occurs. This proposal is based upon the view that sentences are processed from left-to-right and that predictiveness is important in the processing of sentences. That is, words early in a sentence somehow limit the set of alternatives possible later. Then the earlier the serial position in a sentence at which ungrammaticalness is introduced (given the same type of distortion), the greater the interference with predictiveness, the greater the uncertainty concerning what may follow, and, it is hypothesized, the less grammatical the string will be judged to be.

In the experiment reported here, judgments of grammaticalness of semi-grammatical sentences were obtained from fluent native speakers of English. It was expected that semi-sentences with inversions of word-order at the beginning of the string would be judged less grammatical than semi-sentences with equivalent inversions serially later. In addition, it might reasonably be expected that strings with inversions that disrupt the boundaries of major phrases would be judged less grammatical than strings that leave the boundaries of major phrases intact.

METHOD

Materials. Three sets of test-sentences were constructed. The first set (active sentences) was formed as follows. First, a simple sentence was constructed: "The boy hit the ball." This sentence could be converted into another meaningful, grammatical sentence by reversing the Noun-Phrases that appear in the Subject and Object, producing: "The ball hit the boy." Four strings were derived from each of these sentences by the inversion of a proximal pair of words. Both of the original sentences are of the syntactic form Determiner$_1$-Noun$_1$-Verb-Determiner$_2$-Noun$_2$. The semi-sentences derived from them were formed by inversion of the order of the first proximal Determiner and Noun (Determiner$_1$-Noun$_1$, *e.g.* "Boy the hit the ball."), inversion of the second Determiner and Noun (Determiner$_2$-Noun$_2$, *e.g.* "The boy hit ball the."), the first Noun and the proximal Verb (Noun$_1$-Verb, *e.g.* "The hit boy the ball."), and the Verb and the second Determiner (Verb-Determiner$_2$, *e.g.* "The boy the hit ball."). Thus two sub-sets of five strings each, including the undistorted sentences, were constructed.

Another set (passive sentences) was constructed as follows. First, the two undistorted active sentences were transformed into passives: "The ball was hit by the boy." and "The boy was hit by the ball." Each passive thus had a structural form like that of the actives, except that here the Verb is a compound of *was*-Verb-*by*. Four strings were derived from each of the passives in an identical manner to those above, so that again two sub-sets of five strings each, including the undistorted sentences, resulted. An example of a Verb-Determiner$_2$ string derived from a passive sentence would be "The ball the was hit by boy."

The final set (infinitive sentences) was constructed in the following manner. First, a new sentence was formed: "To establish new concepts requires major changes." The reversal of Noun-Phrases in this sentence resulted in another sentence: "To establish major changes requires new concepts." Both sentences have the syntactical form *To*-Verb$_1$-Adjective$_1$-Noun$_1$-Verb$_2$-Adjective$_2$-Noun$_2$. Four strings were derived from each normal sentence by the inversion of the first Adjective and Noun (Adjective$_1$-Noun$_1$) and the second Adjective and Noun (Adjective$_2$-Noun$_2$), and by the inversion of the first Verb and Noun-Phrase (Verb$_1$–Noun–Phrase$_1$) and the second Verb and Noun–Phrase (Verb$_2$–Noun–Phrase$_2$). These inversions produced, for example, "To establish concepts new requires major changes.", "To establish new concepts requires changes major.", "To new concepts establish requires major changes.", and "To establish new concepts major changes requires.", respectively. Thus again two sub-sets of five strings each, including the undistorted sentences, were constructed.

Procedure. For the sub-set of active sentences that contained "The boy hit the ball." and the four strings derived from it, each string was typed on a 5 x 8-in. card, and the set of five cards was presented to 15 *S*s; the order of the cards was shuffled for each *S*. *S* was instructed to rank-order the sentences from the one that seemed to be the best English sentence, which was to be assigned the rank "1," to the one that seemed to be the worst, assigned the rank "5." No additional instructions were given.

For the other sub-set of active sentences and for the two sub-sets of passive sentences and two sub-sets of infinitive sentences, each sub-set was printed on dittoed sheets, along with the instructions. For each sub-set, an order of appearance of the sentences on the sheet was randomly selected and presented to 8 *S*s, the reverse order to 7 others. Thus the five sentences of each sub-set were presented to 15 *S*s. No *S* was presented more than one sub-set. All the *S*s were men and women enrolled at Harvard University.

TABLE I

FREQUENCY THAT EVERY TYPE OF DERIVED STRING WAS JUDGED MORE GRAMMATICAL THAN EVERY OTHER TYPE

ACTIVE SENTENCES
(N=30)

Type of derived string	Determiner$_2$–Noun$_2$	Verb–Determiner$_2$	Determiner$_1$–Noun$_1$	Noun$_1$–Verb
Normal	30‡	30‡	30‡	30‡
Determiner$_2$–Noun$_2$		23†	25†	22*
Verb–Determiner$_2$			20 n.s.	26‡
Determiner$_1$–Noun$_1$				18 n.s.

* $p<0.02$; † $p<0.01$; ‡ $p<0.001$.

Results. The frequency that every type of derived string was rated more grammatical than every other type for the active, passive, and infinitive sentences is contained in Tables I-III, along with the corresponding probability (two-tailed) according to the binomial expansion (sign test). That is, the rank-orders were reduced to pair-comparisons. Since the results for the various types of derived string for the two sub-sets of each set were practically identical, the results for the two sub-sets that comprise each set have been combined.

From Tables I-III, it is clear that for all three syntactic types of sentence (active, passive, and infinitive) the undistorted sentence was judged to be more grammatical than all of the semi-sentences. With regard to semi-sentences with inversions that occurred within Noun–Phrases, it can be seen that in all three cases the earlier in the string the inversion occurred, the less grammatical the string was judged to be. That is, for both active and passive sentences, the strings with inversions of the first Determiner and Noun were judged less grammatical than the strings with inversions of the second Determiner and Noun (though for the passive sentences the difference did not reach statistical significance), and for the infinitive sentences the strings with inversions of the first Adjective and Noun were judged less grammatical than the strings with inversions of the second Adjective and Noun. The difference, however, between the first and second Verb–Noun–Phrase inversions for the infinitive sentences was small, both in the direct comparison and with regard to Adjective$_1$–Noun$_1$, but in the predicted direction.

TABLE II

FREQUENCY THAT EVERY TYPE OF DERIVED STRING WAS JUDGED MORE GRAMMATICAL THAN EVERY OTHER TYPE

PASSIVE SENTENCES
($N=30$)

Type of derived string	Determiner$_2$–Noun$_2$	Determiner$_1$–Noun$_1$	Verb–Determiner$_2$	Noun$_1$–Verb
Normal	30‡	30†	30‡	30‡
Determiner$_2$–Noun$_2$		19 n.s.	22*	30‡
Determiner$_1$–Noun$_1$			16 n.s.	29‡
Verb–Determiner$_2$				30‡

* $p<0.02$; † $p<0.01$; ‡ $p<0.001$.

TABLE III

FREQUENCY THAT EVERY TYPE OF DERIVED STRING WAS JUDGED MORE GRAMMATICAL THAN EVERY OTHER TYPE

INFINITIVE SENTENCES
($N=30$)

Type of derived string	Adjective$_2$–Noun$_2$	Adjective$_1$–Noun$_1$	Verb$_2$-Noun–Phrase$_2$	Verb$_1$-Noun–Phrase$_1$
Normal	30‡	30‡	30‡	30‡
Adjective$_2$–Noun$_2$		22*	25†	25†
Adjective$_1$–Noun$_1$			22*	25†
Verb$_2$–Noun-Phrase$_2$				16 n.s.

* $p<0.02$; † $p<0.01$; ‡ $p<0.001$.

Finally, for both the active and passive sentences, there is a general tendency for the strings with inversions that disrupt major phrase-boundaries (Noun$_1$–Verb and Verb–Determiner$_2$) to have been judged less grammatical than the strings with Determiner–Noun inversions. It might also be noted that for both active and passive sentences Noun$_1$–Verb was judged to be less grammatical than Verb–Determiner$_2$. This might be expected from the serial-position hypothesis, though it is not clear that these two types of inversion can be considered syntactically equivalent.

Discussion. Fluent native speakers seem to be able to judge the relative grammaticalness of semi-sentences (at least the grammaticalness of those types of semi-sentences used here), and inter-*S* agreement appears to be high. This has not always been the case.

> Hill, for example, interpreted the results of some rather casual experiments that he conducted as evidence that the *S*s are not very well able to judge the grammaticalness of sentences.[4] His data, however, do show a moderate degree of inter-*S* agreement. Certainly one can never expect perfect agreement in any empirical undertaking. Thus Maclay and Sleator obtained judgments of grammaticalness, meaningfulness, and ordinariness for six types of verbal material ranging from zero-order approximations to English up to normal sentences, found a moderate level of inter-*S* agreement, and concluded that *S*s can judge grammaticalness.[5] It is interesting to note that even though the questions Maclay and Sleator asked their *S*s—whether the various strings were grammatical, meaningful, or ordinary—were quite different, the three scales that were derived were not very different for the six types of string. In fact, for only two of the six types of material did the ordinal positions vary from one scale to any other by as many as two steps.

The results of this experiment seem to support the hypothesis that the earlier in a string a particular type of word-inversion occurs, the less grammatical the string will be judged to be. This was seen to occur regardless of the syntactic nature of the original sentence, whether active, passive, or infinitive. The only possible exception was the Verb–Noun–Phrase inversion in the infinitive sentences, which showed practically no difference for the two serial positions, though the difference that did occur was in the predicted direction. The use of two original sentences in which Noun–Phrases in the Subject and Object reversed from one to the other makes it unlikely that the results are explicable in terms of the particular words used: for example, the sequence "boy the" appeared as the earlier Determiner–Noun inversion for half of the *S*s and as the later inversion for the other half who received active (and passive) sentences.

The results of this experiment are quite similar to the results of a sentence-identification experiment.[6] The reaction-time to semi-sentences that were constructed in a manner identical to those used in this experiment

---

[4] A. A. Hill, Grammaticality, *Word,* 17, 1961, 1-10.

[5] H. Maclay and M. D. Sleator, Responses to language: Judgments of grammaticalness, *Intern. J. Amer. Ling.,* 26, 1960, 275-282.

[6] L. E. Marks, Some structural and sequential factors in the processing of sentences, *J. verb. Learn., verb. Behav.* (in press).

(where the *S*s had to decide which of two normal sentences was the source of each semi-sentence) conformed quite well to the serial-position hypothesis: the earlier in the string the inversion occurred, the longer the latency of response. In fact, even the Verb–Noun–Phrase inversion in infinitive sentences showed this serial effect.

The effect of serial position may be examined with regard to other variables of which it might be a function. Chomsky has developed the distinction between surface and deep structures in the analysis of sentences.[7] The surface structure of a sentence corresponds to what is phonologically

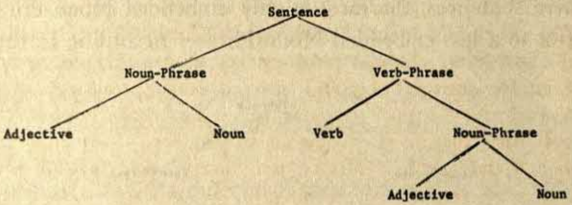


FIG. 1. SURFACE STRUCTURE OF ACTIVE SENTENCES

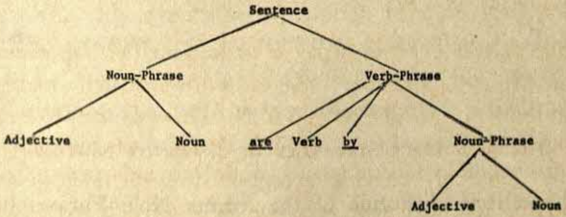


FIG. 2. SURFACE STRUCTURE OF PASSIVE SENTENCES

represented in the sentence; the deep structure corresponds to what is semantically interpretable. The surface structure, and, essentially, the deep structure, for the active sentences is depicted in Fig. 1. When these sentences were passivized, however, a different surface structure resulted (Fig. 2). But the deep structure of the passivized sentences remains identical to that depicted in Fig. 1. Since the major result of this experiment—that the earlier in a sentence a particular distortions appears, the less grammatical the sentence is judged to be—was seen to hold for both active and passive sentences, it may be concluded that the effect of serial position involves the surface structure, not the deep structure, of the sentence. If the deep structure were involved, then the effect should have reversed itself in the passive sentences, since the order of the logical Subject and Object reverses in passive sentences when going from deep to surface structures.

---

[7] Chomsky, *Aspects of the Theory of Syntax*, 1965, 16.

The variable of hierarchical level might be derived from Chomsky's model as an alternative explanatory hypothesis.[8] That is, the level of judged grammaticalness might be expected to be a function of the degree of embedding in the surface structure of the sentence at which the distortion occurs: the higher the hierarchical level, the less grammatical the string. While this explanation is adequate to account for the results for active and passive sentences, it does not account for those for infinitive sentences. As can be seen in Fig. 3, which depicts the surface structure of the infinitive sentences, the most highly embedded Noun-Phrase appears serially prior to a less embedded Noun-Phrase. According to the hierarchical-level hypothesis, distortion of the former Noun-Phrase should result in a string judged to be more grammatical than distortion of the latter Noun-Phrase. But this was not the case: The left-to-right effect was manifested in infinitive sentences also. This is not to deny that hierarchical level may be a significant variable in the determination of levels of judged grammaticalness. In fact, Coleman found a high correlation between ratings of the grammaticalness of semi-sentences and levels of grammaticalness as determined by Chomsky's formal analysis.[9]

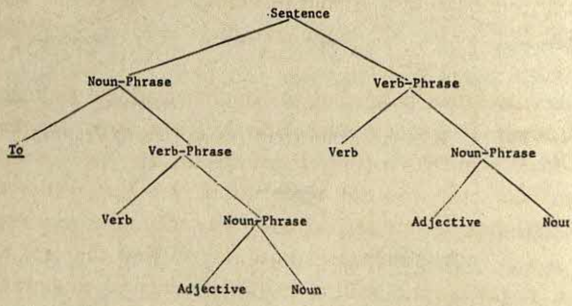


FIG. 3. SURFACE STRUCTURE OF INFINITIVE SENTENCES

The present study, however, did not have as its aim a test of Chomsky's model. Indeed, from his most recent statements it is difficult to determine just what Chomsky's position is with regard to judgments of grammaticalness.[10] For example, though Chomsky states that "there is no way to avoid the traditional assumption that the speaker-hearer's linguistic intuition is the ultimate standard that determines the accuracy of any proposed grammar, linguistic theory, or operational test," he completes the sentence with the qualification that "this tacit knowledge may very well not be immediately available to the user of the language."[11] Further on, Chomsky states that "grammaticalness cannot, in any event, coincide with the intuitive notion of 'deviance' . . ."[12] Perhaps this last point is related to the distinction Chomsky makes between acceptability and grammaticalness. This distinction rests upon the more basic distinction between competence and performance, between the implicit knowledge of the rules of a language and the actual use of the language. Thus Chomsky states that "acceptability is a concept that belongs to the study of performance, whereas grammaticalness belongs to the study of competence," and continues, "Like acceptability, grammaticalness is, no doubt, a matter of degree, but the scales of grammaticalness and acceptability do not coincide."[13]

---

[8] *Ibid.*, 148-152.

[9] E. B. Coleman, Responses to a scale of grammaticalness, *J. verb. Learn. verb. Behav.*, 4, 1965, 521-527.

[10] Chomsky, *op. cit.*, 18-21.

If it is assumed that judgments of grammaticalness do reflect the linguistic intuitions of native speaker-hearers, then these judgments would be the ultimate standard for any linguistic theory. But, Chomsky claims, these judgments only concern acceptability, so, apparently, this would not be so. The precise ontological status that Chomsky gives to linguistic intuitions is far from clear. One possibility is that they are merely constructs in a formal system, with no possible empirical verification. According to this formulation, sentences would be grammatical if they are generated by a grammar; but obviously, if not modified, this formulation of grammaticalness would be utterly trivial. On the other hand, Chomsky may mean that linguistic intuitions can be tapped by some operational test, though not directly. The difficulty that this possibility poses is in knowing when one has reached the linguistic intuitions of a native speaker, since the latter are not directly available to him. In fact, it would appear that the only way that one could know when one has reached these linguistic intuitions would be if one knew in advance what they were.

Perhaps the best solution is to accept the native speaker's judgments of grammaticalness as the ultimate standard for a grammar. The major point of this study was to demonstrate that judgments of grammaticalness are highly dependent upon the psychological operations that occur when verbal material is processed. It is claimed, by virtue of the fact that spoken sentences are processed from left-to-right and written sentences are typically read that way, that words in the first part of a sentence make words later on more predictable, and that the earlier in a sentence a distortion appears, the greater the interference with predictability and processing and the less grammatical the string will be judged to be. In other words, the explanation given here for the results of this experiment is in terms of factors related to performance. If this is so, then perhaps the distinction between competence and performance should be given up. This would have the unfortunate effect, however, of removing the distinction between some factors in language behavior that do seem intrinsically different (for example, syntactic rules versus limitations due to memory). At very best, however, it seems that the distinction between competence and performance should be looked upon as a useful heuristic device that will probably be discarded when some advanced level of psycholinguistic knowledge is reached.

---

[11] *Ibid.*, 21.
[12] *Ibid.*, 158.
[13] *Ibid.*, 11.

SUMMARY

Fluent native speakers of English rank-ordered, according to their grammaticalness, normal sentences and semi-grammatical sentences that were derived from the normal sentences by the inversion of a pair of proximal words or a word and a proximal phrase. The results were taken as support for the hypothesis that judgments of grammaticalness are dependent upon the ways in which people process linguistic materials. It was found that, regardless of the syntactic nature of the sentence investigated, for semi-sentences with syntactically equivalent types of inversion, the earlier in the string the inversion occurred, the less grammatical the string was judged to be. This result was predicted on the basis of the hypothesis that left-to-right predictiveness is important in the processing of sentences. It was also found that, in general, semi-sentences in which the inversions of word-order disrupted the boundaries of major phrases were judged to be less grammatical than semi-sentences in which the inversions left the boundaries of phrases intact. The results of the experiment were examined in the light of some formal linguistic hypotheses concerning semi-sentences and with respect to Chomsky's general theoretical formulation.

PERCEPTUAL LEARNING RESULTING FROM EXPOSURE TO A STIMULUS-INVARIANT

By WILLIAM EPSTEIN, University of Kansas

The visual angle subtended by an object does not specify the distance of the object. A *pair* of different retinal subtenses does, however, produce a reliable and predictable perception of relative distance.[1] The same is true of continuous sequence of size-transformations.[2] Thus, although a single retinal subtense is ineffective, a difference between visual angles associated with simultaneous or successively presented stimuli effectively controls perceived distance. Simultaneous or successive size-transformations of visual angle provide information about distance.

The present experiment was designed to study the effects of a procedure planned to enhance the effectiveness of the single retinal subtense. The training procedure took advantage of the demonstrated effectiveness of size-transformations of visual angle. On a pretest, *O* judged the distance of a standard that was presented with all distance-cues eliminated. Then the *O*s in the experimental condition were exposed to sequences of continuous size-transformations associated with movement of the standard in *O*'s saggital plane. At regular intervals during the movement, accurate and precise information about the distance of the standard was provided. In other words, *O* was exposed to transformations in retinal size that were correlated with precise information about changes in distance. In the post-test, *O* judged the distance of the stationary, isolated standard again. (Two control conditions will be described later.)

The following reasoning led to the choice of the training procedure: A single retinal subtense does not specify distance because it is compatible with a practically infinite number of combinations of physical size and physical distance. This ambiguity might be eliminated if *O* has the oppor-

---

* Received for publication September 24, 1966. This research was supported by Grant MH 4225 from the United States Public Health Service. The article was written while the author was a senior Postdoctoral Fellow of the National Science Foundation.

[1] William Epstein and S. N. Baratz, Relative size in isolation as a stimulus for relative perceived size, *J. exp. Psychol.*, 67, 1964, 507–513; William Epstein and Samuel Franklin, Some conditions of the effect of relative size on perceived relative distance, this JOURNAL, 78, 1965, 466–470.

[2] W. H. Ittelson, Size as a cue to distance: Radial motion, this JOURNAL, 64, 1951, 188–202; William Schiff, Perception of impending collision: A study of visually directed avoidance behavior, *Psychol. Monogr.*, 79, 1965, Whole No. 604, 3-18ff.

tunity to learn that the changes in retinal size for the given standard are uniquely correlated with changes of physical distance. The training procedure provides this opportunity. If *O* can learn the function that relates transformations of retinal size to distance, then, when he is confronted again by the isolated retinal subtense, the distance of the standard should be specified by the size of the subtense. On the post-test, *O* should be able to make fairly accurate judgments of distance despite the absence of the conventional distance-cues.

Method: (1) General plan. A 3 × 2 factorial experiment was performed, with three conditions of training, two standard distances, and two standard sizes. The 60 *O*s were distributed evenly among the 12 conditions.

(2) Observers. The *O*s were undergraduates with no prior experience in experiments on perception. They were assigned to the experimental conditions in groups of 12.

(3) Standards. The standards were 3-in. and 5-in. squares cut from black cardboard and covered with fluorescent paint.

Viewing-box. An oblong viewing-box, 10 ft. long and 21 in. in height and width, was constructed. The interior surfaces were painted flat black. On the floor of the box, at the base of the left-hand wall, four 23-in. ultraviolet tubes were placed. These tubes provided almost constant illumination of the fluorescent standard at all locations in the box. On the left-hand wall of the box, seven small light-boxes were mounted. They were located at 1-ft. intervals beginning at 3-ft. from the face of the viewing-box. The face of each light-box was covered by black cardboard from which a number had been cut, designating the distance of the light-box from the face of the viewing-box. A track was laid along the length of the floor. Microswitches were placed at 1-ft. intervals along the track. Each light-box was wired to a corresponding microswitch in such a way that a standard traversing the track would activate each microswitch in succession and illuminate the corresponding light-box for a fraction of a second, thereby revealing the distance of the standard at that moment.

A cart was constructed upon which the standard could be mounted. The cart was so connected to a pulley-system that *O* could move the cart back and forth along the track by turning a crank in the manner of a fishing reel. The movement of the cart was accompanied by some noise, but the noise was effectively masked by the sound of an electric fan.

Measuring rod. Size-judgments were obtained with the help of a stainless-steel bar (51.5 × 1.2 cm.) on which two black aluminum sleeves (each 10.2 × 2.5 cm.) were fitted.[3] The sleeves were fitted to slide easily across the bar, but with sufficient friction that they would not move of their own accord. The *O*s reported their estimates of the size of the standard by adjusting the separation between the inside-ends of the two sleeves. The maximal separation allowed by the device was 11.8 in., which proved to be more than sufficient for all measurements.

Procedure. The first stage was a practice-session in which size-judgments were obtained for a set of 10 squares ranging in size from 1–10 in. The squares were located either 3, 6, or 9 ft. from *O*. This session was conducted in a room other than that used for the experimental testing. The practice-judgments were made under full-cue conditions, but without visual inspection of the measuring rod. In other words, the *O*s provided tactual matches of their visual impressions of size.

---

[3] This device was first used by Hans Wallach, M. E. Moore, and Linda Davidson, Modification of stereoscopic depth-perception, this JOURNAL, 76, 1963, 191–204.

Immediately following the practice session, *O* was led, with his eyes closed, into the completely dark experimental room and positioned at the face of the viewing-box. Throughout the experiment, *O*'s head was held stationary by a chin-rest and head-vice arrangement. To eliminate the monocular distance-cues, *O* was restricted to monocular vision through a 4-mm. aperture that was covered by a red filter. Under these conditions of viewing, *O* could see only the fluorescent standard and the number illuminated by the light-box when the corresponding microswitch was activated.

There were three main conditions: (1) In the experimental condition (*Ex*) *O* began by judging the size and distance of one of the standards at either 4 or 8 ft. The size-judgments were made tactually by using the measuring device. The distance-judgments were verbal estimates. Then *O* moved the standard back and forth along the track between the 3-ft. and 9-ft. distances. Five movement-cycles were allowed, followed by a 30-sec. interval of rest and then a second session of five cycles. Prior to the initiation of the movement, *O* was informed that the same square would be used throughout the experiment, and the meaning of the illuminated numbers was explained. After the moving exposure, *O* repeated the original size- and distance-judgments. (2) and (3) Two control conditions were used. One (C_1 Transformation only) differed from the Experimental Condition (*Ex*) in only one way; the main power-source turned off that the light-boxes could not be illuminated. This condition provided exposure to the same sequence of size-transformations presented in Condition *Ex*, but without accompanying information about the distance of the standard. In the other (C_2, Scale only), the standard was masked during the training period but the power to the light-boxes remained. This condition provided exposure to the distance-scale presented in Condition *Ex* but no exposure to transformations of size.

Each *O* in each condition judged only one standard (3 or 5 in.) at only one distance (4 or 8 ft.). The control-conditions (C_1 and C_2) were introduced to determine whether changes observed between the two tests in Condition *Ex* might be due solely to exposure to size-transformations (C_1), or opportunity to learn a distance-scale (C_2). It was expected that only under Condition *Ex* would *O* learn to use the single retinal subtense as a basis for judging distance since only this condition provided information about size-distance correlations.

Results: (1) Practice. The practice-data were not analyzed statistically. Visual inspection of the data showed that the *O*s could reproduce visually perceived extents tactually with little error and very small inter-*O* variability.

(2) Distance-judgments. Table I shows the mean distance-judgments for the 12 conditions of the experiment. The pretest means ranged from 1.7–3.3 ft. for the 4-ft. distance, and from 1.6–4.6 ft. for the 8-ft. distance. Analysis of variance of the pretest judgments showed that no

TABLE I

MEAN DISTANCE-JUDGMENTS (FT.) BEFORE (*B*) AND AFTER (*A*) THREE CONDITIONS OF TRAINING

Condition	Distance												Combined diff.
	4 ft.						8 ft.						
	3-in standard			5-in standard			3-in standard			5-in standard			
	B	*A*	diff.	*B*	*A*	diff.	*B*	*A*	diff.	*B*	*A*	diff.	
Ex	1.7	4.8	3.1	3.3	4.5	1.2	3.3	6.1	2.8	3.9	7.7	3.8	2.73
C_1	2.3	2.4	0.1	2.8	2.6	−0.2	2.7	2.2	−0.5	4.6	5.8	1.2	0.15
C_2	3.3	4.4	1.1	1.7	2.6	0.9	1.6	3.1	1.5	2.9	3.8	0.9	1.10
Combined	2.4	—	—	2.6	—	—	2.5	—	—	3.8	—	—	—

TABLE II

MEAN SIZE-JUDGMENTS (IN.) BEFORE (*B*) AND AFTER (*A*) THREE CONDITIONS OF TRAINING

Condition	Size												Combined diff.
	3 in.						5 in.						
	4-ft. distance			8-ft. distance			4-ft. distance			8-ft. distance			
	B	*A*	diff.	*B*	*A*	diff.	*B*	*A*	diff.	*B*	*A*	diff.	
E	2.27	2.85	0.58	1.84	3.40	1.56	3.43	5.10	1.67	2.87	3.80	0.93	1.19
C_1	3.00	3.40	0.40	1.80	2.72	0.92	3.43	4.30	0.87	2.52	4.23	1.71	0.98
C_2	2.67	3.47	0.80	1.18	1.98	0.80	1.72	2.71	0.99	2.60	2.73	0.13	0.68
Combined	2.65	—	—	1.61	—	—	2.87	—	—	2.66	—	—	—

significant differences associated either with treatment, size of the standard, or distance of the standard. The fact that the 4-ft. and 8-ft. distances were not discriminated indicates that the experimental arrangement did indeed eliminate the cues for distance that are normally available.

Table I shows all the mean distance-judgments obtained after the training and the differences between the pre- and post-training judgments. The difference-scores were transformed by the addition of a constant to eliminate negative scores and an analysis of variance was performed. Only the main effect of training conditions was significant ($F = 6.94$, $df = 2/48$, $p < 0.01$). The combined means in the right-hand column of Table I reveal the nature of the main effect. The difference-scores were greater under Condition *Ex* than under Conditions C_1 and C_2. Furthermore, *t*-tests showed that only Condition *Ex* yielded a mean difference-score significantly different from zero.

The distance-estimates under Condition *Ex* increased significantly and became more accurate. Prior to the exposure to the distance-correlated size-transformations, the mean estimates were 2.5 and 3.6 ft. for the 4- and 8-ft. distances, respectively. After exposure, the mean judgments were 4.65 ($SD = 1.39$) and 6.90 ($SD = 1.96$) for the two distances. The analysis of variance reported above has shown that the two estimates obtained *prior* to the training exposure did not differ. A *t*-test was used to evaluate the difference between the post-exposure judgments of the 4- and 8-ft. distance, and a *t*-value of 9.20 was obtained with 18 *df* ($p < 0.01$), which means that after exposure to distance-correlated size-transformations the *O*s could discriminate two distances that were indiscriminable before the exposure.

(3) Size-judgments. Table II shows the mean size-judgments before and after training. Prior to training, the mean judgments ranged from 1.18–3.00 and from 1.72–3.43 for the 3-in. and 5-in. standards, respectively. An analysis of variance of the pretraining judgments showed no significant differences were associated either with treatment, distance, or size of the standard. Both standards were underestimated at both distances, and the 3- and 5-in. standards were not distinguished. Although the mean size-judgments tended to be smaller at 8 ft. than at 4 ft., the difference was not significant. As another check of this latter finding, the Mann-Whitney *U*-Test was used to compare the size-judgments for the 4- and 8-ft. distances ($N = 30$ in each distribution).[4] The value of z

---

[4] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences,* 1956, 116–127, esp. 124. The *U*-test is one of the most powerful nonparametric tests. Although it is not entirely appropriate to follow a nonsignificant analysis of variance by a nonparametric test, the procedure was adopted here because the size-distance relationship being tested is a very well established one.

(corrected for ties) was 2.358, which permits rejection of the null hypothesis ($p < 0.01$ with a one-tailed test). These results are compatible with the common finding that, in the absence of distance-cues, size-estimates are inversely related to distance.[5]

An analysis of variance of the difference-scores (transformed by the addition of a constant to eliminate the signs used to distinguish between increases and decreases from the pre- to the post-test) yielded no significant effects. It would appear that all the training conditions were equally effective in modifying the estimates of size. The *t*-tests computed for the difference-scores for each training-condition, separately, showed that the mean difference-score for each condition differed significantly from zero: Condition *Ex*, 1.19 ($SD = 1.70$), $t = 3.236$ ($df = 19$), $p < 0.01$; Condition C_1, 0.98 ($SD = 1.89$), $t = 2.28$ ($df = 19$), $p < 0.05$; Condition C_2, 0.67 ($SD = 1.02$), $t = 2.50$ ($df = 19$), $p < 0.05$.

(4) Perception of motion. During the observation of motion, the *O*s were encouraged to describe the appearance of the standard. In evaluating these reports, it should be remembered that *O* had been informed explicitly that the same standard square would be used throughout the experiment. Furthermore, the movement of the standard was produced by *O* himself. Under Condition *Ex* ($N = 20$), there were: (a) 12 reports that the standard appeared constant in size while it moved back and forth; (b) 5 reports that the standard contracted and expanded as it moved back and forth; (c) 1 report that the standard always remained constant in size and stationary while *O* felt himself to move; and (d) 2 reports that the appearance of the standard alternated between Descriptions *a* and *b*. Under Condition C_1 ($N = 20$), there 12 reports of (a), five reports of (b), one report of (c), and two reports that (d) the standard was stationary and only changing in size.

It might be expected that the effectiveness of the training-procedure would depend at least in part on the way the moving standard appeared to *O*. To explore this possibility, the mean difference-score for the *O*s who gave Description (a) was compared with the mean for all others combined. The mean difference-score for distance under Condition *Ex* was 2.70 ($SD = 1.74$) for the 12 *O*s who gave Description (a) and 2.13 ($SD = 2.20$) for the 8 *O*s who gave the remaining descriptions. The mean difference-score under Condition C_1 was 0.20 for the 12 *O*s who gave description

---

[5] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21–37; William Epstein, John Park, and Albert Casey, The current status of the size-distance hypotheses, *Psychol. Bull.*, 58, 1961, 491–514.

(a) and 0.12 for the remaining 8 *O*s.[6] The differences between the subgroups in each condition were not significant. It should be noted, however, that the critical variable of experience is probably the perception of movement, and in this regard the main group, *i.e.* Description (a), and the smaller group did not differ greatly.[7]

Discussion. The chief objective of the present experiment was to determine whether *O* could extract an invariant from a continuous transformation, and whether this newly acquired relationship would modify his responses to an otherwise ambiguous static stimulus. The invariant in question was the ratio of visual angle to physical distance. The ambiguous static stimulus was the retinal subtense corresponding to a single isolated standard. The primary variable was the judgment of distance. The results show that exposure to a sequence of visual angle-distance correlations (Condition *Ex*) does significantly modify distance-perception. The absence of significant difference-scores under Condition C_1 indicates that exposure to transformations of retinal size without the accompaniment of information about distance does not affect perceived distance. Nor is exposure to a distance-scale in the absence of size-transformations sufficient to modify perception (Condition C_2). The singular effectiveness of Condition *Ex* was expected, since only this condition provided the opportunity for extracting the size-distance invariant. The ineffectiveness of Condition C_1 also is understandable. Continuous symmetrical transformations of retinal size can produce reliable impression of radial motion, *i.e.* changes of *relative* distance, but these transformations do not specify the *absolute* distance of the standard at any given point in the sequence. The ineffectiveness of Condition C_2 is important to our analysis only as it minimizes the possibility that the effectiveness of Condition *Ex* is simply the result of acquainting *O* with an appropriate range and scale of distance.

In addition to the primary response-variable, a secondary response-variable was measured. The selection of judged size as a response-variable was natural, since size- and distance-perception often are considered to be interdependent. The results presented in Table II show that all the training conditions modified the judgments. There was no statistical evidence that the conditions were differentially effective. This finding was unexpected, particularly in the case of Condition C_2. Speculation about the meaning of this finding is deferred until it can be substantiated in subsequent investigations.

---

[6] There were too few *O*s to permit finer individual comparisons between the subgroups.

[7] The same analysis was carried out for the difference-scores for size. Again, there was no evidence that those scores depended on the appearance of the moving standard.

What is the nature of the learning that occurred under Condition *Ex*? It is best described in terms of Gibson and Gibson's differentiation-theory of perceptual learning which stresses learning about stimuli rather than the learning of responses.[8] During the moving exposure, *O* learned about a higher order stimulus-variable—the size-distance invariant—which transmuted the originally ineffective potential stimulus into an effective stimulus. If this description is correct, then this experiment has contributed to the meagre store of empirical demonstrations of the type of perceptual learning that is proposed by the theory.[9]

---

[8] J. J. Gibson and E. J. Gibson, Perceptual learning: Differentiation or enrichment?, *Psychol. Rev.*, 62, 1955, 33–40.

[9] J. J. Gibson, Perception as a function of stimulation, in Sigmund Koch (ed.), *Psychology: A Study of a Science,* Vol. 1, 1958, 456-501. William Epstein and John Park, Examination of Gibson's psychophysical hypothesis, *Psychol. Bull.,* 62, 1964, 180-196.

GLOBAL ERRORS IN CHILDREN'S REPRODUCTIONS OF PROGRESSIVELY CHANGING PICTURES

By MELVYN SCHNALL, Brandeis University

The present research deals with the way children organize visual events which involve progressive changes in objects over time. In the simplest case of such an event, a single figure or object may be depicted as progressively changing in position, size or some other attribute. The focus of this research, however, is on the more complex type of event in which two objects are shown doing different things in relation to each other at different times.

Two studies have analyzed developmental features of spatiotemporal integration and provide the background for the present research. One study by Olum dealt with developmental differences in the perception of causality.[1] She employed a stimulus display of the sort developed by Michotte in which a black rectangle moves to a red rectangle and may give the impression of hitting the red out of the display field.[2] The presence or absence and the quality of the causal impression among adults depends on the ratio of the speeds of movement of the two rectangles.[3] The main finding of Olum's research with seven year olds in comparison to adults was that the children reported phenomena never or very rarely experienced by adults. These included "passing" phenomena, in which the rectangles are reported to pass over each other, and "mutual approach" phenomena, in which activity is reported to originate in *both* of the rectangles such that they hit, bump, or push *each other.* Olum suggests quite plausibly that the "mutual approach" may be the result of stroboscopically induced movement in the red rectangle due to special features of the display. Regarding the "passing" phenomena, she suggests the possibility that a child is relatively less capable of breaking down the configuration provided by the two movements in space and time. Can we, however, be more specific about the sorts of cognitive processes which the adult employs in construing the spatiotemporal configuration and which the child fails to employ?

The second study forming the background of the present research is pertinent to this issue.[4] Instead of employing materials which provide adequate stimulation

---

* Received for publication July 25, 1966. This investigation was supported by Public Health Service Research Grant No. HD-01931 from the National Institute of Child Health and Human Development. The author is grateful to Mr. Paul Lashua, Principal of Oak View School, Bloomfield, N.J., and his teachers for their cooperation in making subjects available. The assistance of Stephen Bendich, Joel Hetler, and Alan Waldman is gratefully acknowledged.

[1] Vivian Olum, Developmental differences in the perception of causality, this JOURNAL, 69, 1956, 417-425.

[2] Albert Michotte, *The Perception of Causality,* 1963, 29-31.

[3] Michotte, *op. cit.,* 41-252.

[4] Melvyn Schnall, Spatio-temporal integration in progressively changing visual

for real or apparent movement perception, the present author used series of discrete pictures which portrayed intermittent steps in the progressive change of position or size of two figures. The figures were a red and a blue disk or a red and a blue stick-figure. The *S*s of different ages were shown the series and then asked to draw what they had seen in the order presented and also to tell in words what they had seen. It was found that with increased age, the *S*s increasingly predicate qualities of motion and change when these are not perceptually present, but are rather integrative notions applied to discrete stimuli.

Of immediate interest here is the fact that some of the children drew and described simultaneous changes in both the figures when only one figure at a time was shown undergoing change in the stimulus-series. For example, instead of one figure moving toward another, both were shown moving simultaneously side by side, or toward each other. Furthermore, the children's drawings sometimes showed a figure undergoing an appropriate change in the context of the stimulus-series, but it was the wrong color.

These "errors" appeared remarkably similar to the sorts of "passing" and "mutual approach" phenomena which had been reported by Olum's younger *S*s. In this author's study, however, there was no perceived movement involved and no reasonable way in which stroboscopically induced motion could, even if it were present, account for the findings. The inappropriate description of change in a non-changing object was viewed as a failure in the differential attribution of different qualities to different objects in a complex event. The cognitive structurization by the child may thus be termed global as opposed to articulated in the terms of Heinz Werner's developmental approach to cognition.[5]

The study to be reported here involves an attempt to replicate the finding of global errors in reproductions of discrete pictures showing events having two figures, and to analyze the effect of (a) same vs. different coloration of the figures; and (b) presence vs. absence of an "anchoring" feature in the stimulus field. Regarding same vs. different coloration of figures, we are interested in the effects that differential attribution of color has on the task of differentially attributing properties of progressive change and stability. The anchoring feature is designed as a potentially meaningful, contextual basis for maintaining stability in one figure while the other is undergoing change in position. Thus, one set of stimulus conditions touches upon the level of figural attributes and the other upon the level of contextual meaning.

Method: (1) Subjects. One hundred twenty eight children were studied in a small elementary school serving a predominantly middle-class neighborhood. The sample contained 32 children in grades 2, 3, 5, and 6. Within each grade level, four groups

---

patterns, microfilmed Unpublished Doctoral dissertation, Clark University, 1964, 38-48; Spatio-temporal integration in progressively changing visual patterns, in Seymour Wapner and Bernard Kaplan, (eds.), *Hein Werner, 1890-1964,* 1966, 63-70.

[5] Heinz Werner, *The Comparative Psychology of Mental Development,* Rev. ed., 1957, 104-142; The concept of development from a comparative and organismic point of view, in D. B. Harris, (ed.), *The Concept of Development,* 1957, 125-148.

were established which were matched closely for age, sex, and school ability as judged by the teacher and principal. Each group was shown a different variation of two series (Series A and Series B) with order of presentation controlled.

(2) Stimuli. The stimulus-series consisted of five pictures drawn on 5 x 8-in. file cards with a blank inserted between each picture. The pictures and blanks were each exposed for approximately 2 sec. through a small opening in a plywood board. Both series were designed to depict two figures undergoing some change together, followed by one of the figures remaining in a changed or displaced state while the other returned to its original state.[6]

Series A: The field consists of a 2-in. square within which, on the bottom line, are two small circles, one on top of the other. Successive pictures show the bottom circle elongating, lifting the top circle until it is almost touching the top line of the inside of the square. The top circle remains in this displaced position while the bottom figure returns to its original size and shape.

Series B: The field consists of a 3½-in. horizontal line. In the beginning, two small circles are side by side on the right-hand end of the line. Successive pictures show them moving to the other end of the line (always touching). The circle on the left remains displaced at the left-hand end of the line while the circle on the right returns to its original position on the right-hand end of the line.

Given the basic outlines of Series A and B, the following variations were introduced in order to study the effects of same vs. different coloration of figures and presence vs. absence of an anchoring feature. (See Fig. 1.)

Variation I: Color of circles different (red and blue), anchor present. (Sereis A: Anchor consists of two hooked lines extending down from the top edge of the square providing a 'receptacle' for the top circle; Series B: Anchor consists of a 'pocket' or semicircular ditch at the left-hand end of the horizontal line into which the left circle falls.)

Variation II: Color of circles different, no anchor.

Variation III: Color of circles same (both red), anchor present.

Variation IV: Color of circles same, no anchor.

(3) Procedure. After seeing each series, *S* was asked to draw the pictures in the same order in which he had seen them on separate sheets of paper on which were printed all the field features appropriate to the series and variation which *S* had been shown. *S* had merely to draw in the circles with appropriately colored pencils. After the drawings were completed, *S* was asked to tell in words what he had seen in the pictures, and his verbal responses were tape-recorded.

(4) Method of analysis. Detailed analyses were performed on the drawings revealing a variety of types of errors. These may be simplified for purposes of analysis into three exhaustive categories.

(a) Correct. This category includes drawings which are perfect or which are essentially correct (*i.e.* extra steps drawn in depicting progressive changes or some step omitted).

(b) Global errors. A global error involves a failure to articulate which specific things are undergoing which specific changes at which specific times, while never-

---

[6] The general design of the series was selected to maximize the possibility of *S*'s depicting simultaneous changes in the two figures, specifically, the displaced figure returning to its original state along with the other figure, based on preliminary findings by Schnall, *op. cit.,* 1964, 38-48, 95-100.

theless manifesting an appreciation of the kinds of changes portrayed in the stimulus series. The major form of global error is the complete or partial return of the displaced object to its original position at the end of the series. Other forms of globality may be present however, including the depiction of the displaced object

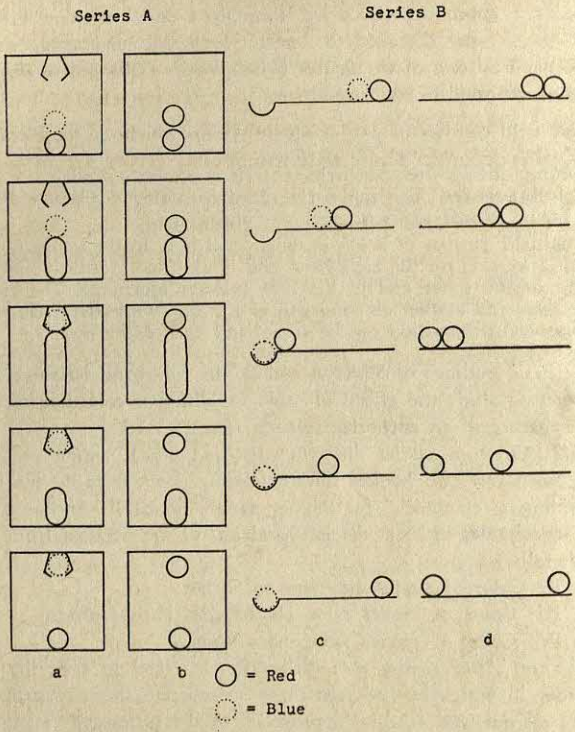


FIG. 1. EXAMPLES OF VARIATIONS IN THE TWO STIMULUS-SERIES
(Variation I: color of figures different, anchor present shown in a and c; Variatoin IV: color of figures same, no anchor shown in b and d.)

in the displaced position throughout the entire series. A few cases meet the general definition of a Global error but are specific to one or the other series.

(c) Other errors. Included in this category are, for the most part, sets of drawings in which absolutely nothing systematic could be detected. These are, thus, unscorable from the point of view of distinguishing systematic errors. Included also are instances which reflect profound failure to articulate the spatial features of the pictures, instances where systematic alternation of color of one figure and the other is present, and instances in which systematic changes are present but so unrelated to the changes depicted in the stimulus series that they cannot be considered global or correct.

Results. It was found that 20% of the drawings fell in the global category on both Series A and B, demonstrating that this type of error may be found with considerable frequency in the reproduction of temporally extended events in which no real or apparent motion perception is present.

Movement and "change" when reported by the *S*s in the present study have the status of a conceptual integration of discrete materials, *i.e. S* imputes motion to the figures in the series of discrete pictures. Integrated verbal descriptions involve attributing progressive change to the figures in the series by means of active verbs such as "to go," "to get," "to roll," "to grow." In order to show that the global errors are associated with such integrative notions, an analysis was performed on the relation between drawing score and presence vs. absence of integration as revealed in the verbal descriptions. Table I shows that the global errors differ significantly from the other errors; the 'Other errors' are not associated with verbal integration while the 'Global errors' are (Series A, $X^2 = 14.29$, $p < 0.001$; Series B, $X^2 = 7.29$, $p < 0.01$). The 'Global score' does not differ from the 'Correct score,' both being associated with verbal integration (Series A, $X^2 = 1.15$; Series B, $X^2 = 1.55$, $0.20 < p < 0.30$ in both cases).

TABLE I

RELATIONSHIP BETWEEN DRAWING SCORE AND VERBAL INTEGRATION
(Number of *S*s)

Series	Drawing score	Verbal integration	
		present	absent
A	Global	18	8
	Other error	13	39
	Correct	40	10
B	Global	14	10
	Other error	8	26
	Correct	50	20

Regarding age-differences (see Table II), younger *S*s make more errors of all sorts than older *S*s, but significant age-differences regarding the 'Global errors' specifically are not manifested here. Table II also shows that Series B was easier to reproduce correctly than Series A.

Table III contains breakdowns of the data for analysis of the effects of anchor and color differentiation. In comparisons of the anchor vs. no anchor conditions (I and III vs. II and IV), the numbers of *S*s making global errors do not differ on either Series A or Series B. It appears that the presence vs. absence of an anchoring feature in the field has no effect on globality. On Series A, however, the 'Other errors,' in comparison to

'Correct' and 'Global' combined, are positively and significantly related to the presence of the anchor ($X^2 = 6.32$, $df = 1$, $p < 0.05$ two-tailed). It appears that the highly unfamiliar anchoring contraption in Series A was generally disruptive, but not specifically related to the spatio-temporal integration of the series.

Same vs. different coloration of the figures does, however, appear to have a specific effect upon the global errors. A comparison of the 'Global' vs. 'non-Global' scores in relation to conditions of color different vs. same (I and II vs. III and IV) is significant on Series B ($X^2 = 5.12$, $df = 1$, $p < 0.05$ two-tailed). The same analysis for Series A is not significant ($X^2 = 3.08$, $df = 1$, $0.10 < p < 0.20$ two-tailed), but is in the same direction as the Series B analysis. The *S*s in the 'Other error' category are equally distributed between color same vs. different. The data may be viewed as supporting the conclusion that color differentiation produces increased globality.

TABLE II

DRAWING SCORES IN DIFFERENT SCHOOL GRADES AND STIMULUS-SERIES
(Number of *S*s)

Series	Drawing score	School grade		Total
		2, 3	5, 6	
A	Global	14	12	26
	Other error	35	17	52
	Correct	15	35	50
B	Global	15	9	24
	Other error	27	7	34
	Correct	22	48	70

TABLE III

EFFECTS OF ANCHOR VS. NO ANCHOR, AND COLOR DIFFERTENT VS. SAME
(Number of *S*s)

Series	Variation	Drawing score		
		Global	Other	Correct
A	Anchor	14	33	17
	No anchor	12	19	33
B	Anchor	10	16	38
	No anchor	14	18	32
A	Color diff.	17	26	21
	Color same	9	26	29
B	Color diff.	17	18	29
	Color same	7	16	41

Discussion. The major findings of the present research are: (a) 'Global type' errors may be observed in children's drawings of visual events involving more than one object under conditions which rule out real or apparent movement-perception; (b) 'Global type' errors differ from 'Other drawing errors' in that the global are significantly related to presence of integrative, active-verb usage in the verbal descriptions, whereas the 'Other errors' are associated with no verbal integration; (c) 'Global errors' are made more frequently when the figures are differently colored than when they are the same color. "Anchoring" field features appear to have no specific effect upon globality.

The problem for *S* in viewing any complex event is not only to detect what changes are present, but to subordinate differentially these changes to conceptions of particular objects or substances. Movement and progressive change must be treated cognitively as figural attributes. In adult experience, there is decisive and virtually immediate subordination of attributes to object concepts; objects dominate experience, with the state of their properties, stable or changing, always subordinate to them. In contrast, the 'Global errors' suggest that children sustain a far looser relationship between substantive and attributive, or object and property aspects of experience. It is as though the children note changes predominantly and only secondarily subordinate these to conceptions of particular objects. The result is frequently a systematic set of changes, but, in relation to the physical stimuli, there is a failure to articulate which specific things are undergoing which specific changes at which specific times.

This formulation is similar to Werner's description of global conception among children, which characteristically involves the relegation of particular constituent elements or objects to some dominant quality of the total situation.[7] Where attributive changes dominate experience and override constituent objects, one observes *simultaneous* change, '*mutual* approach' and other 'Global errors.' In light of this, we may interpret the effect of color differences as having complicated the task for the children since the two colors, as figural attributes, had also to be differentially subordinated to different objects. Since the 'anchoring' field feature is less involved at the level of figural attributes, it has no specific effect upon the appearance of 'Global errors.'

SUMMARY

Children have been observed to report simultaneous changes in both objects of stimulus-events which show only one object changing at a time,

---

[7] See Footnote 5.

i.e. 'Global type' errors. In the present study, children were shown series of pictures, one picture at a time, depicting progressive changes in two objects. Effects on 'Global type' errors of same vs. different color of objects and presence vs. absence of an 'anchoring' field feature were studied. It was found that: (1) 'Global type' errors occur in children's drawings of visual events under conditions which rule out real or apparent motion-perception; (2) 'Global type' errors differ from other drawing errors in that the global are associated with integrative, active-verb usage in the verbal descriptions of the stimulus series, whereas the 'Other errors' are not; (3) 'Global type' errors are made more frequently when the objects are differently colored than when they are the same color. 'Anchoring' field features appear to have no specific effect upon globality.

THE RECALL OF LONG AND SHORT STIMULUS-LISTS

By JOHN CERASO, Yeshiva University

An experimental finding which has not received much attention in recent years is the following: If a long list is learned to the same criterion as a short list, delayed recall will be better for the long list.[1] Why should this be true? One might think, *a priori,* that intralist interference would lead to *more* forgetting for the long list. A possible answer is that, despite the fact that the *criterion* of learning is the same for both, the long list is better learned. One might reason as follows: Learning-scores are recall-scores obtained during the learning period; at every learning point, more interference is generated in the long list than in the short list; to surpass their respective 'interference barriers' and reach a given criterion, the items of the long list require more strength than those of the short list; and the greater strength reveals itself in better delayed-recall. This analysis is consistent with the fact that it takes more trials to learn a long than a short list. Other evidence for this interpretation is presented in this study.

EXPERIMENT I

Procedure. Two groups of 16 *S*s were used in this experiment. The first group, the 10*L* Group, learned a list of 10 paired associates, while the second group, the 20*L* Group, learned a list of 20 paired associates. Learning was terminated for both groups after the first trial on which at least 60% of the items were correctly anticipated. The material was presented on a memory-drum at a 2-sec. rate, with 4 sec. between successive groups of 10 pairs. Three random orders of each list were used in presenting the material. A delayed recall was given one week after learning. Each *S* was given a sheet containing all stimulus-items in alphabetical order and asked to write the appropriate response next to each stimulus. There was no time-limit to this task.

Materials. The stimuli were 20 nonsense-syllables taken from Glaze's list.[2] Their associative values ranged from 40-60%. The responses were 20 one-syllable nouns, with a Thorndike-Lorge frequency of 5-17. All 20 pairs were learned by the 20*L*

---

* Received for publication October 5, 1966. This research was supported by Grant GB 1604 from the National Science Foundation and by funds from the Research Committee of the Ferkauf Graduate School of Yeshiva University. The author is indebted to Ann Henderson and Joel Goldberg for assistance in collecting the data.

[1] R. S. Woodworth, *Experimental Psychology,* 1938, 60; J. A. McGeoch, *The Psychology of Human Learning,* 1942, 374-375.

[2] J. A. Glaze, The association value of non-sense syllables, *J. genet. Psychol.,* 35, 1928, 255-267.

Group. For the 10*L* Group, two sets of 10 pairs each were used, half the *S*s learning each set.

Subjects. The *S*s were students at New York University. They were randomly assigned to the groups on the basis of order of appearance at the laboratory.

Results. The data on learning and recall summarized in Table I are in agreement with previous results. To reach the same criterion as the 10*L*-*S*s, the 20*L*-*S*s required more trials and they demonstrated better recall. Both effects are statistically significant ($p < 0.01$).

One might, however, question whether the items of the short and long lists reach the same criterion of learning. Easy items in the long list might be learned just as rapidly as in the short list, the greater number of trials being required by the more difficult items. The easy items would then receive a large number of overlearning trials and reach a higher criterion than comparable items in the short list. Greater recall for the long list could then be attributed to the easy items which are overlearned.

TABLE I

LEARNING AND RECALL IN EXPERIMENT I

Group	Mean trials	Criterion-trial Total correct	 % correct	Delayed recall Total correct	 % correct
10*L*	7.44	103	64.38	60	37.50
20*L*	13.31	203	63.44	167	52.19

To deal with this possibility, an item-analysis of the data was performed. The point was to determine: (1) if the easy items of the long list did indeed reach higher criterion-levels than did short-list items; and (2) if the long-list superiority in recall was limited to the easy items. Items were ranked in five classes according to the total number of anticipations achieved prior to the criterion-trial. Since the ranking was done for each *S*, there were two items in each class for the 10*L*-*S*s and four items in each class for the 20*L*-*S*s. It was then possible to determine for each rank the probability of a correct response on the criterion-trial (the total given divided by the total possible). The probabilty of delayed-recall also was computed (the total given on the delayed-recall trial divided by the total given on the criterion-trial).

These values are presented in Table II. There is some tendency for the easier items of the long list to have higher recall-probabilities on the criterion-trial, but the difference is slight and does not seem large enough to explain the substantial superiority of long-list items in *all* classes.

It is important to realize that items of the same rank in 10*L*- and 20*L*-

groups are equated in terms of ease of material, *S*s represented (all *S*s are represented equally for each rank), and (roughly) for probability of occurrence on the criterion-trial. The fact that these items show better delayed recall for the 20*L*-Group seems a strong argument for the position that there is a higher degree of learning for the 20*L*-Group.

A final analysis of the data is based on a definition of equal learning in terms of the absolute number of anticipations per item. If to be anticipated a given number of times, items in the long list must be better learned, items anticipated equally often should show better delayed recall in the long as compared with the short list. Since the maximal number of times an item could be correctly anticipated by the 10*L*-Group was 6, the analysis is confined to items receiving 0-6 anticipations. All items achieving a given number of anticipations prior to the criterion-trial were pooled and the percentage of these items given on the learning trial was obtained. The probability of recall on the delayed test also was obtained by dividing the number of items recalled by the number learned. These data are presented in Table III. While the scores for 10*L*-items are slightly higher on the criterion-trial, the scores for the 20*L*-items are better in delayed recall.

TABLE II

Probability of Recall (%) on the Criterion-Trial and the Delayed-Recall Trial as a Function of Item-Rank in Experiment I

Rank	Criterion-trial		Delayed recall	
	10*L*	20*L*	10*L*	20*L*
1	45.31	25.06	51.93	94.85
2	42.81	42.92	40.36	77.19
3	57.28	63.25	62.36	67.41
4	82.81	88.89	57.25	79.33
5	95.31	97.14	66.66	91.96

TABLE III

Probability of Recall (%) on the Criterion-Trial and the Delayed-Recall Trial as a Function of Number of Anticipations in Experiment I

Number of Anticipations	Criterion-trial		Delayed recall	
	10*L*	20*L*	10*L*	20*L*
0	44.62	27.08	48.28	92.31
1	53.85	59.26	57.14	84.37
2	77.42	80.00	62.50	50.00
3	89.47	61.29	52.94	100.00
4	100.00	92.86	33.33	76.92
5	100.00	94.73	88.89	94.44
6	100.00	92.86	100.00	76.92

The results of this analysis favor the hypothesis that items equally well anticipated are better learned in the long list. It must be borne in mind in interpreting these results that an analysis confined to items of 0-6 anticipations omits from consideration the easiest items for the long list (roughly the top 25%) and therefore underestimates the superiority in recall for the 20*L*-Group.

EXPERIMENT II

In Experiment I, the final criterion-trial was a paced trial, while the delayed-recall trial was unpaced. One might argue that retention over the recall-interval of one week was not better for the *S*s with the long-list, but that it was the shift from paced to unpaced test which favored the long-list *S*s. This argument implies that it is harder to recall from the long than from the short list *only when recall is paced.* Is it also true that it is harder to recall from the long list when recall is unpaced? To answer this question, one must have an unpaced criterion-trial on which *S*s with long- and short-list reach the same criterion of learning.

TABLE IV

LEARNING AND RECALL IN EXPERIMENT II

Group	Criterion-trial		Delayed recall		% retained
	Total correct	% correct	Total correct	% correct	
10*L*	195	65.00	66	22.00	31.70
20*L*	337	56.17	160	26.67	43.17

Method. Because of the difficulty of predicting from *a prior,* paced trial what score would be achieved on the unpaced criterion-trial, the fixed-trails method was used. Preliminary testing indicated that five trials for the short list and seven trials for the long list would yield a mean criterion-score of 60% learned on the unpaced trials.

Experiment II, then, proceeded as follows: Learning trials were conducted in the same way as in Experiment I. After the fifth trial for the short list, and the seventh trial for the long list, *S* was given a sheet containing all the stimuli and asked to write the appropriate response next to each stimulus. He was given as much time as he wished to complete the recall sheet. Another unpaced recall-test was given one week later. The 60 *S*s—30 in each group—were recruited from various colleges in Manhattan and were tested at Yeshiva University.

Results. The data are organized in the same way as in Experiment I. Table IV summarizes the over-all results for learning and recall. Note that the long-list *S*s fell short of the 60% desired criterion-level, while the short-list *S*s exceeded it, although the difference between the groups was not significant. The scores of delayed-recall show a difference in the opposite direction with the long list *S*s showing better recall, although again the difference was not significant. There is evidence, however, for the superiority of long-list recall, which is to be found in Table IV under "percentage retained." The data given under this heading represent the mean percentage of recall-scores over learning-scores. The long-list *S*s are significantly superior on this measure ($p < 0.05$). Even with achievement measured on an unpaced criterion-trial, the long-list *S*s show better retention.

Table V shows that, while the items of the short list were consistently better recalled on the criterion-trial, the long-list items were consistently better retained over the one-week interval. Table VI presents the data on number of anticipations. As in the comparable Table III for Experiment I, the data are somewhat mixed, although it seems fair to say that the short-list items do somewhat better on the criterion-trial but somewhat more poorly over the interval. Again, it should be noted that the easiest 20*L*-items are omitted from consideration (the top 10%). Thus, the outcomes of the three analyses of the Experiment II resemble the outcomes of Experiment I.

TABLE V

PROBABILITY OF RECALL (%) ON THE CRITERION- AND THE DELAYED-RECALL TRIAL AS A FUNCTION OF ITEM-RANK IN EXPERIMENT II

Rank	Criterion-trial		Delayed recall	
	10*L*	20*L*	10*L*	20*L*
1	47.55	30.00	17.62	39.31
2	50.14	36.70	22.73	30.77
3	61.41	52.35	23.08	43.39
4	80.36	68.94	39.60	48.62
5	91.38	85.28	48.58	70.71

TABLE VI

PROBABILITY OF RECALL (%) ON THE CRITERION- AND THE DELAYED-RECALL TRIALS AS A FUNCTION OF NUMBER OF ANTICIPATIONS IN EXPERIMENT II

Number of Anticipations	Criterion-trial		Delayed recall	
	10*L*	20*L*	10*L*	20*L*
0	41.27	20.55	17.31	18.33
1	65.00	79.49	23.08	35.48
2	93.18	92.19	26.83	40.68
3	97.30	87.23	69.44	53.66
4	100.00	93.49	43.48	76.74

Discussion. When two groups learn lists of verbal material at different rates, one of three methods can be employed to equate the groups for degree of learning. One method is to have both groups learn an equal number of items (or, when the lists differ in length, to have both learn the same percentage of items, as in the present study). Another technique is to compare items from the two groups which have reached the same probability of occurrence on the criterion-trial. A third method is to compare items which have been correctly anticipated the same number of times. In the present study, 10*L*- and 20*L*-Groups were compared for delayed recall after being equated for learning by each of these methods. In each case, the results suggest better recall for the 20*L*-Group. Two ways of interpreting these data are possible. One might argue that the long and short lists were indeed equated for degree of learning but that list-length *per se* influences recall. Another possibility is considered in this paper; namely, that, in order to reach a given criterion, long-list items must be better learned (because of greater intralist interference) and this better learning reveals itself in the slower rate of forgetting for the long-list group. Since there seem to be no good reasons for accepting the first position, and some very good reasons (mainly derived from interference-theory) for rejecting it, the second interpretation seems more acceptable. There are two implications of the data, one methodological and one theoretical, which derive from this analysis.

The methodological point is quite simple. There is a large number of interesting problems in retention which lead one to compare the retention of groups that learn at different rates. To mention a few: A list of high similarity is learned more slowly than a list of low similarity; the second of two lists is learned at a different rate from the control list with which it is compared in studies of proactive inhibition; and so forth. In studies of this kind, groups learning at different rates have been taken to the same criterion of learning. Underwood, for example, has developed a number of techniques to insure criteria of equal learning,[3] but taking groups to equal *learning criteria* is not necessarily the same as taking them to an equal degree of learning. Conclusions about retention which have been based upon equating groups by a criterion-method must, therefore, be reëxamined in cases where there is legitimate doubt as to whether a criterion-measure is adequate to assess learning strength.

An example of such a case may be Underwood's work on the problem of whether people who learn quickly retain better or more poorly than people who learn slowly.[4] Underwood, using a procedure which involved comparing the recall of items from fast and slow learners which reached the same learning criterion, concluded that fast and slow learners do not differ in retentive capacity. It may be, however, that items of slow learners have to be *stronger* than items of fast learners to reach an equal learning criterion. If this were the case, equivalent delayed-recall levels would indicate poorer retention for the slow learners. To put it differently, Underwood may have succeeded in showing that a strong item for slow learners has the same slope of retention as does a weaker item for fast learners. At least part of the difference between fast and slow learners may be a differential sensitivity to interference, requiring higher degrees of learning by slow learners to reach equivalent criterial levels as fast learners. Some support for this notion is to be found in the fact that the ratio of the mean number of trials to criterion for the 8 fastest 20*L*-*S*s of Experiment I to the mean number of trials to criterion for the 8 fastest 10*L*-*S*s was 1.75. The same ratio computed for the 8 slowest learners is 2.01. These ratios suggest that slow learners suffer more from increased length of list than do fast learners. For Experiment II, the ratios were 1.41 for slow learners and 1.04 for fast learners. Again slow learners seem to be affected more strongly by list-length than are fast learners.

---

[3] B. J. Underwood, Degree of learning and the measurement of forgetting, *J. verb. Learn. verb. Behav.*, 3, 1964, 112-129.

The second point to be made is a speculative one which may be useful in understanding *RI* and *PI*. Ceraso, Schiffman, and Becker have argued that in a two-list interference experiment both first and second lists show retention-loss with the passage of time.[5] A possible explanation of this loss is the following: Immediately after learning the second list, if the *S* is instructed to "give me the first (or second) list-responses to this stimulus," he may be able to limit his search for the correct response to the appropriate list. That is, the instruction to recall from a given list may allow a selective mechanism to operate.[6] It may be, however, that one of the effects of time is on the operation of the selective mechanism. If that were the case, *S* would in effect be recalling from a list of greater length (more items) than the control group. If the facts presented here are correct, it would follow that experimental *S*s should recall more poorly than control *S*s, because experimental *S*s have learned the material to a degree necessary to perform in a short list, but not to the degree necessary for performance in a long list. Thus, delayed *RI* and *PI* may be understood as reflecting the same kind of interference which occurs during learning.

SUMMARY

If a long list is taken to the same criterion of learning as a short list, a delayed-recall test shows better recall of the long list. In reaching the same criterion, the longer list will require more trials. These facts suggest that the long list has been learned to a greater degree than the short list when the criterion of learning is reached. A methodological implication of this argument is that when the effect of some variable on recall is studied, and when that variable also affects the rate of learning, it is not satisfactory to equate groups for degree of learning by bringing them to the same criterion. A theoretical implication is that delayed *RI* and *PI* may be a manifestation of the same type of interference which operates during learning.

---

[4] Underwood, Speed of learning and amount retained: A consideration of methodology, *Psychol.* Bull., 51, 1954, 276-282.

[5] John Ceraso, Dave Schiffman, and Benson Becker, Recall interference in retroactive inhibition, *J. Psychol.*, 59, 1965, 259-265.

[6] B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Learning,* 1960

SIZE-ESTIMATION OF FAMILIAR OBJECTS UNDER INFORMATIVE AND REDUCED CONDITIONS OF VIEWING

By H. RICHARD SCHIFFMAN, Rutgers University

When a viewer attempts to judge some characteristics of a familiar object, he has two kinds of information available. There is information given by current visual cues, and also information as a result of past experience with similar objects. The traditional emphasis upon immediate visual cues in estimating the size of familiar objects has been recently questioned. For example, Bolles and Bailey obtained visual, and non-visual estimates from memory, of the sizes of familiar objects uniform with respect to size.[1] Correlations of both the non-visual and visual estimates with the measured sizes of the objects were high (of the order of 0.99) and the authors concluded that information given by past experience with familiar objects is a primary determinant of their judged size, assuming that the object is correctly identified. It can be argued, however, that the similarity of the size-estimations in the visual and non-visual conditions does not necessarily indicate that an identical process is determining the response.

It is possible that size estimates made from memory are accurate for familiar objects and that estimates made with the use of visual cues may be equally accurate. This does not necessarily imply that a memorial process operates when the object is visually presented. The determinants of apparent size may depend in a given case upon the information available. When judging size in a visually informative environment, it appears reasonable that *S* makes use of the visual information directly rather than using this information to identify the object and then referring to past experience with similar objects for the size-judgment. In contrast, when judging in a visually impoverished environment, the identification of the object and the memory of its size may become influential in determining the judgment.

---

* Received for publication September 30, 1965. Based upon a Doctoral dissertation submitted to the University of North Carolina and directed by Dr. Samuel Fillenbaum.

[1] R. C. Bolles and D. N. Bailey, Importance of object recognition in size constancy, *J. exp. Psychol.*, 51, 1956, 222-225; see also, A. C. McKennell, Visual size and familiar size: Individual differences, *Brit. J. Psychol.*, 51, 1960, 27-35.

Slack reported evidence that knowledge of size was a cue to apparent size.[2] His experiment examined the effect of familiar size under conditions providing many gradients of texture like those described by Gibson.[3] Chairs that were too large and too small were viewed in a large, flat field at various distances. The *S*s judged the size of the small chair to be larger and that of the large chair to be smaller than the sizes of unfamiliar control objects, that is, there was a regression in size estimates toward known familiar size. Slack concluded from these results that familiar size should properly be considered as one of the important cues to apparent size.

Fillenbaum, Schiffman, and Butcher conducted a replication and extension of Slack's study which used additional control objects and also recorded distance estimates to permit an examination of the relationship between apparent size and apparent distance.[4] Slack's results were not replicated. The findings of this second experiment gave no indication of a systematic bias either in size or distance estimates, as a function of familiar size. It is possible that the results of Fillenbaum, Schiffman, and Butcher may have been due to use of exaggerated differences in the size of the familiar objects. These objects might have suggested too obviously to *S* that he was viewing specially contrived stimuli. This information could have reduced or eliminated the effect of familiar size.

There is little direct evidence to resolve this dilemma, to say what are the determinants of apparent size under normal circumstances. The present study has as its general hypothesis the following: (1) when visual information on the size of familiar objects is ample, size estimates will be primarily determined by this information; (2) as visual information of the size of familiar objects is reduced, without the loss of proper identification, size-estimates will be more dependent upon information given by past experience with similiar objects.

The stimuli for the present experiment were moderate enlargements and reductions of familiar objects coming from classes that are relatively uniform in size. Comparison-objects were both familiar objects unaltered in size and unfamiliar objects exactly equal in size to the off-sized experimental objects. Three conditions of viewing, offering differing amounts of visual information, were used.

Method: (1) Subjects. Sixty-three men from an introductory course in psychology served as the *S*s. All were naïve as to the content of the experiment.

(2) Apparatus. The apparatus allowed *S*, looking from a stationary position, to estimate the size of different objects under the three different viewing conditions.

---

[2] C. W. Slack, Familiar size as a cue to size in the presence of conflicting cues, *J. exp. Psychol.*, 52, 1956, 194-198.

[3] J. J. Gibson, *The Perception of the Visual World*, 1950, 66-71.

[4] S. Fillenbaum, H. R. Schiffman, and James Butcher, The perception of off-size versions of a familiar object under conditions of rich information, *J. exp. Psychol.*, 69, 1965, 298-303.

Stimuli and viewing conditions could be changed between trials without *S*'s knowledge, and all responses by *S* could be timed.

S was seated in a backed stool facing a window at eye level about 5 in. from him. The window was opened and closed by *E*. When the window was closed, a sheet of opaque material was placed in front of the window. The window was 12 in. high by 10½ in. wide. A 25-w. bulb mounted above *S* remained on at all times.

The objects to be judged could be seen through the open window mounted on screens constructed of bleached canvas stretched over plywood. The *near* screen was 39½ in. wide by 4 ft. long and mounted 26 in. from the floor. It was located 4 ft. from *S*'s window. In the two *far* conditions, the same screen was used. It was 64 in. wide by 4 ft. long and was 26 in. from the floor. It was 26 ft. from the window. The *far-reduced* condition made use of a series of reduction-tubes. Their lengths and diameters were so varied that the object was seen monocularly surrounded by about a 2-in. border of canvas.

The walls and floor of the room behind the window had textured surfaces. The lighting was supplied by a series of flood lamps mounted near the floor and ceiling on the side of the window not facing *S* and directed toward the screens. The lighting was so controlled by *E* that there was slightly under 2 ft.-c. incident light at the near screen and slightly above 2 ft.-c. incident light at the far screen.

The stimuli were six normal-sized objects, six off-sized objects made to look like the normal objects, six unfamiliar objects, and two wooden rectangles. The normal-sized objects were a Lark cigarette package, a playing card, a roll of toilet paper, an examination-booklet, a manila folder, and the cover of a *Life magazine* (February 21, 1964). Of the off-sized, the three smaller (*Lark* package, toilet paper roll, and manila folder) were reduced 15% from normal size; and the three larger (playing card, examination-booklet, and *Life* cover) were enlarged 15% from normal size. The original colors in all cases were preserved.[5]

Each unfamiliar abstract construction (control object) was made from a rectangular surface with various three-dimensional shapes, fashioned from cardboard, attached to its surface. It was painted flat black. To facilitate the comparison with the off-sized objects, these constructions were made the same height and width as the objects.

The wooden rectangles were painted flat black; one 9 in. high by 2 in. wide; the other 4 in. high by 2 in. wide.

(3) Procedure. Each *S* was tested individually. He was informed that he was performing an experiment in size-estimation and was instructed to estimate the height of the objects shown as quickly and as accurately as possible. Sizes of the objects were to be reported to the nearest ¼ in.

Three practice trials were given with a neutral stimulus. After the practice trials all the *S*s were presented with the 4 in. and 9 in. sticks of wood. These stimuli were presented in the same order for all the *S*s and were shown at all viewing conditions. Following presentation of the wood stimuli, the *S*s were presented with the experimental objects.

(4) Design. This experiment may be regarded as six replications of one basic experiment, one for each set of experimental objects used. Each replication would

---

[5] A subsidiary study performed prior to the one reported here indicated that 15% off-sizings were discriminable from normal sizes of the objects.

involve the comparison of three objects: a normal-sized object, a reduced or enlarged version of this object, and an unfamiliar construction having the size of the off-sized object. Each object was judged at each viewing condition, near, far, and far-reduced. It was assumed that a bias in judgment might result if *S* saw any object more than once either in its normal size or off-size, so it was decided to limit each *S* to one of the three objects from each replication. Accordingly, each *S* was randomly assigned to one of the nine cells (3 objects by 3 viewing conditions) of each replication. Since there were six replications, each *S* saw six objects, each object coming from the set of three related objects that made up a given replication. Hence, all observations in a given replication represent size estimations of different *S*s, and are independent of each other.

Results. Analyses of variance and covariance were performed on each of the six replications comprising the size-estimates of three different objects estimated under three viewing conditions. The covariant measure was a ratio based on the size estimates made by all the *S*s of the 4 in. and 9 in. wood stimuli at the near and far viewing conditions. The covariance analyses were used to adjust all estimates of the experimental objects to a common metric; however, these analyses did not add significantly to the results to the analyses of variance, and it is the latter analyses and unadjusted scores that are here reported. The analysis of variance performed on the scores for all objects and all viewing conditions yielded a within-cells estimate of error, based on data in all cells, and this was used to make individual comparisons between pairs of means specified by the predictions.

Six comparisons were made for each replication, using the mean estimates. The normal-sized object was compared with the off-sized object, and the off-sized object with the control object, at each of the three viewing conditions. A difference between normal-sized and off-sized objects indicates the use of visual information in the size estimates. A difference between the off-sized and control objects, with the off-sized object approaching the size of the normal object, indicates the influence of a memorial process on the estimates.

The comparisons between the estimates at all viewing conditions are given in Table I. Column A presents the *t*-values for the comparisons between the normal-sized and off-sized objects for all replications. Column B gives the corresponding values for the comparison of the off-sized object to the unfamiliar control object.

The results for the *near*-viewing condition indicates that estimates of the off-sized objects approximate those of the control objects; none of the *t*-values is significant. On the other hand, estimates of the off-sized objects significantly differed from those of the normal objects in four out of six replications. For all replications the estimates of the off-sized objects differed from those of the normal objects in the direction of the size of the control objects. These results indicate the use of visual information on estimates of the off-sized objects in the near condition.

Column A for estimates in the *far*-viewing condition indicates that the estimates of the off-sized objects are not statisically different from those of the normal objects for four out of six replications. The two comparisons where significant *t*-values resulted are from replications where anomalous results were found (off-sized examination-booklet and normal-sized manila folder). The estimates of the off-sized examination-booklet and normal-sized manila folder were greater at the *far* condition than the estimates made of these objects at the *near* and *far-reduced* viewing condition, where the degree of visual information is the greatest and least respectively. These deviant results do not appear so much to run counter to prediction as to be anomalous results precluding any comparisons with predictions. On the other hand, comparisons of the estimates of the off-size and con-

TABLE I

MEANS OF SIZE ESTIMATES AT ALL VIEWING CONDITIONS OF NORMAL-SIZE OBJECTS, OFF-SIZE OBJECTS, AND CONTROL OBJECTS, AND *t* VALUES OF COMPARISONS BETWEEN MEAN ESTIMATES OF OFF-SIZE AND NORMAL-SIZE OBJECTS AND BETWEEN OFF-SIZE AND CONTROL OBJECTS

Replication	Means of normal-sized objects (in.)	Means of off-sized objects (in.)	Means of control objects (in.)	A *t*	B *t*
Near:					
lark package	3.89	2.82	2.50	2.38*	<1
playing card	3.14	4.18	3.93	2.74†	<1
toilet roll	5.11	3.89	3.79	2.65†	<1
examination-booklet	7.71	9.36	10.21	1.51	<1
manila folder	12.14	9.21	10.07	3.22†	<1
Life cover	13.29	15.36	15.21	1.55	<1
Far:					
lark package	3.93	3.53	2.54	<1	2.20*
playing card	3.68	3.93	4.57	<1	1.68*
toilet roll	4.75	4.75	4.00	<1	1.63
examination booklet	8.93	12.00	9.86	2.82†	1.96*
manila folder	13.29	11.46	10.46	2.01*	1.10
Life cover	14.57	13.96	16.29	<1	1.74*
Far-reduced					
lark package	3.86	3.21	2.64	1.44	1.27
playing card	4.14	3.79	3.14	<1	1.71*
toilet roll	5.00	4.43	3.18	1.24	2.74†
examination booklet	8.00	7.64	10.61	<1	1.86*
manila folder	12.03	11.50	5.71	<1	6.36†
Life cover	12.71	14.36	12.14	1.23	1.66

A: *t*-values for comparisons between means of normal-sized and off-sized objects.
B: *t*-values for comparisons between means of off-size and control objects.
* $p<0.05$ † $p<0.01$

trol objects suggest a trend for estimates of the off-sized objects to differ from those of the control objects in the direction of the sizes of the normal objects. This occurred in five out of the five comparisons for replications without anomalous cell mean estimates with *t*-values for three replications reaching statistical significance. In the *far* condition memorial process seems to be operating.

For the *far-reduced* viewing-condition, Column A indicates that the estimates of the normal-sized objects were not significantly different from those of the off-sized objects for all replications. In four out of six comparisons for the estimates of the off-sized and control objects, the *t*-values were significant and for four replications the estimates of the off-sized objects differed from those of the control objects, in the direction of the size of the normal objects. The results, then, for both the *far* and *far-reduced* conditions, indicate a memory influence on the estimates of the off-sized objects.

Discussion. The size-estimates made in the viewing condition offering ample visual information, that is, the *near* condition, emphasize the importance of the visual input beyond the function of object identification. The results for the *far* and *far-reduced* viewing conditions were generally consistent with the prediction that a memorial process would be used when size-estimates were made of objects under relatively reduced visual conditions. The findings imply that if the sizes of familiar objects are known to *S*, this information may be relied on when visual information is reduced; not that visual cues are used only to the extent that memorial evidence is not available under any conditions of viewing as Bolles and Bailey and others appear to argue.[6]

The present experiment, to some extent, presented *S* with the kind of visual information specified by Gibson. In the *near* and *far* viewing-conditions the walls and floor were visible to *S* and presented a grained surface. The finding of a memory effect for familiar objects in the reduced viewing condition is not, however, readily explained by Gibson-type gradients alone. The estimates of the unfamiliar objects in the *near* and *far* viewing-conditions testifies to the sufficiency of visual cues to size when *S* has no other information available about the objects. When, however, the object is a familiar one and is presented in a visually reduced situation, prior knowledge of its size enters as a size-determinant much in the way that Bolles and Bailey suggest. In arguing against the importance attributed to past experience by theorists such as the Transactionalists, perhaps Gibson has neglected the possibility that familiar objects present different information to *S* than unfamiliar objects. The present study has indicated that whether or not this information becomes influential depends largely upon the viewing conditions of the object relative to *S*.

---

[6] Bolles and Bailey, *op. cit.,* 222-225; McKennell, *op. cit.,* 27-35.

SUMMARY

An experiment was performed to determine the processes used in estimating the sizes of different kinds of objects under a viewing-condition offering (1) an ample visual information, (2) a reduced condition, and (3) a monocular reduced condition. The results using familiar objects, off-sized versions of familiar objects, and unfamiliar objects, indicated that, with ample visual information, known size was not one of the operative determinants of estimated size, but a memorial process was used in conditions of reduced visual information. The results suggest that the process used in size-estimation of familiar objects is dependent upon the visual information available.

CORRELATION OF PERFORMANCE IN DETECTING VISUAL AND AUDITORY SIGNALS

By WALTER J. GUNN and MICHEL LOEB,
Fort Knox, Kentucky

The experiments to be described are to test the degree of correlation in detecting signals that are presented in different sensory modalities. Relatively few people have investigated such relationships, although there have been some experiments concerned with the proficiency of *S*s in long-term monitoring (vigilance) tasks, in which the signals to be detected were presented through different sensory channels. In one of these, low coefficients of correlation ($r < 0.30$) were obtained between individual performance on auditory and visual tasks requiring vigilance.[1]

In a recent experiment, Eijkman and Vendrik determined fluctuations over time of an *S*'s sensitivity for small changes in the intensity or in the duration of auditory and visual signals.[2] They found that the fluctuations in detecting increments of intensity were not correlated, while the fluctuations in detecting changes in duration were highly correlated. They concluded that there is probably a 'duration center' common to both vision and hearing, but no corresponding 'intensity center.'

There could be a factor influencing both channels which, however, did not vary in its effects over time, or whose variations were negligible compared to other factors concerned with discrimination of intensity. Its existence could not necessarily be inferred by Eijkman and Vendrik's procedure. Instead it might be necessary to determine sensitivity to intensity change for a number of individuals, and to determine whether an index of sensitivity, such as d', was correlated for the two modalities, vision and hearing.

Even if sensitivity is not correlated across modalities, one might expect there would be personality variables that would produce similarity of

---

[*] Received for publication April 26, 1966. From the U.S. Army Medical Research Laboratory. The authors are indebted to Messrs. Carl E. Guthrie, Ralph J. Garner, Barry L. Zimmerman, and John P. Hamm for technical assistance.

[1] D. E. Broadbent and Margaret Gregory, Vigilance considered as a statistical decision, *Brit. J. Psychol.*, 54, 1963, 309-323; D. N. Buckner, A. Harabedian, and J. J. McGrath, A study of individual differences in vigilance performance, Human Factors Problems in ASW, Tech. Rep. *No. 2, Human Factors Research*, 1960.

[2] E. Eijkman and A. J. H. Vendrik, Can a sensory system be specified by its internal noise? *J. acous. Soc. Amer.*, 37, 1965, 1102-1109.

responding in different detection-tasks. For example, if a person adopted a liberal criterion for responding, he would detect more signals and also make more false responses than another person with the same sensitivity but who adopted a more conservative criterion. In the language of detection-theory, the second person would be characterized by the same d' but a higher β.[3] It would be reasonable that a person adopting a liberal criterion in one modality will do so when he responds in another modality, but this remains to be demonstrated. In the present experiments the *S*s were asked to detect an occasional more intense signal in a train of visual or auditory pulses. The parameters of detection, β and d', were determined from the results for the two modalities, and the coefficient of correlation was calculated for these scores.

EXPERIMENT I

Procedure. Twenty volunteer *O*s, civilian and military personnel assigned to the US Army Medical Research Laboratory, were employed in the experiment. While no rigorous screening procedure was employed, the individuals who served had no obvious auditory or visual deficits.

The experiment lasted somewhat less than an hour for each *O*, and consisted of an auditory and a visual session. A 10-min. rest-period separated the sessions, and the order of the sessions was counterbalanced.

Auditory session. Each *O* listened to a series of 240 pulses of white noise, 0.5 sec. in duration, and 2.5 sec. apart. The pulses of sound were presented through earphones in a moderately quiet environment. Normally the pulses were approximately 70 db. in intensity, but at random intervals the intensity was increased by 1.1 db. These more intense pulses were termed signals and *O*'s task was to respond to these by pressing a key. Responses within 2.0 sec. after the end of the signal were scored as detections; other responses were scored as false responses.

Visual session. Each *O* sat in a semi-reclining position in a field dental chair with his head in a headrest approximately 6 ft. from a source of light. Normally, the room was dark. In the test-session, the light was switched on for 0.5 sec. 240 times at 2.5 sec. intervals.The standard pulse was produced by a single shaded 100-w. bulb. The illumination at the position of *O*'s head was about 3 ft.-candles. On 40 of the 240 pulses, randomly determined, the pulse was produced by two 100-w. bulbs. These more intense flashes, 3 db. greater in intensity, were termed signal pulses. *O*'s eyes were taped shut with transparent plastic tape to eliminate eyeblink and minimize the importance of coupling—or observing response—and to make the visual task more comparable to the auditory one. The brightness of the light that *O* saw is obviously unknown, but the signal pulses presumably were 3 db. more intense than non-signal pulses.

O's task, when he detected a signal pulse, was to press the key. Detections and false responses were scored as they were in the auditory session.

---

[3] J. A. Swets (ed.), *Signal Detection and Recognition by Human Observers*, 1964, ch. 1.

Each *O* was run through an entire sequence of 240 pulses in a practice-period before each experimental period. During this practice, the *O*s were told when they missed a signal or made false responses, but they were not corrected in the experimental sessions.

Results and discussion. The number of signals correctly reported, and the number of false responses were tabulated for each condition for each *O*. From these the normal deviates and ordinates of the normal curve were obtained, and the parameters reflecting sensitivity (d') and conservatism in responding (β) were computed.[4] Table I shows the values obtained. Since there was reason to believe that these values, especially for β, may have changed as a result of experience in the experiment, the values are tabulated separately for those *O*s who detected auditory signals first and for those *O*s who received the visual signals first.

It is apparent that both the βs and d's are larger for the visual conditions. Medians are given because the distributions, especially those for the βs, deviate from normal. Differences in d' between visual and auditory signals were statistically significant ($p < 0.01$, by Wilcoxon *T*-test); the auditory signals were significantly harder to detect. Differences in β were not significant. No significant effects of order were noted for d's, but β was significantly higher for those first detecting visual signals ($p < 0.10$ for visual, $p < 0.05$ for auditory, $p < 0.01$ for both kinds of signals combined, as tested by the Mann-Whitney *U*-test).

TABLE I

MEDIAN β AND d' FOR THE AUDITORY AND VISUAL TASKS IN EXPERIMENT I

Task	β			d'		
	Total group	A-V*	V-A	Total group	A-V	V-A
Auditory	3.02	2.47	5.14	1.86	1.51	2.03
Visual	3.83	2.88	6.98	2.56	2.54	2.47

* A-V is median of those first hearing auditory signals; V-A is median of those first seeing visual signals.

Coefficients of correlation for β and d' across modalities are given in Table II. Because of the nature of the distributions, both Spearman rank order correlations and Pearson product-moment correlations were computed; the Pearson correlations for β were computed after a normalizing log transformation.

Interestingly, there are not only moderate and statistically significant cross-modality correlations for β, presumably reflecting a consistent bias in the criterion for responding, but also comparably large and statistically significant correlations for d'. This latter finding was unexpected. It is particularly interesting in that it may signify, perhaps, a central cognitive component of signal detection which trancends the particular modality and yet is not attitudinal in nature.

---

[4] J. A. Swets and S. T. Sewall, Invariance of signal detectability over stages of practice and levels of motivation, *J. exp. Psychol.*, 66, 1963, 120-126.

Before accepting such a finding and speculating on its possibilities, it seemed necessary to replicate it. A second similar experiment was therefore devised.

EXPERIMENT II

Experiment II differed from the first experiment in two principal respects: (1) an attempt was made to equate more closely the difficulty of the visual and auditory signals; and (2) each *O* detected both auditory and visual signals for two sessions in order to determine the extent to which unreliability attenuated the cross-modality relationship, and to determine more adequately the effect of experience.

TABLE II

CROSS-MODALITY CORRELATIONS OF DETECTION PARAMETERS IN EXPERIMENT I

	β	d'
r_s	.60	.67
r_p	.52	.48

(r_s and r_p are Spearman rank-order and Pearson product-moment correlations, respectively. βs were transformed to their common logarithms before Pearson correlations were computed.)

Coefficients of correlation >0.44 are significant at the 0.05 level; >0.56 at the 0.001 level.

Procedure. The same procedures were followed, generally, as in Experiment I. The auditory signals, however, were made 1.5 db. more intense than the non-signal pulses, rather than 1.1 db. The visual signal was produced by adding a 60-w. bulb to the 100-w. bulb that produced the non-signal pulses of illumination. The intensity of illumination for the non-signal pulses remained 3.0 ft.-candles at *O*'s head position while the intensity of the signal pulses was 4.9 ft.-candles. In other words, the increase in illumination during signal pulses was approximately 2.1 db.

As in Experiment I, order of administration of auditory and visual sessions was counterbalanced across *O*s. Each *O* was given 10 min. of practice, with knowledge of his results, in detecting auditory signals, and 10 min. of practice in detecting visual signals, before actual testing began. The procedure was repeated for each *O* on a subsequent day, employing the same order of administration.

Fourteen enlisted men assigned to the US Army Medical Research Laboratory were employed as *O*s. None had a score below 90 on the Armed Forces Qualification Test; none had an obvious severe auditory or visual deficit.

Results. Median βs and d's are given in Table III. It will be seen that

d' was significantly lower for auditory than for visual signals in the first test day ($p < 0.01$ as tested by the Wilcoxon T) but not for the second. Sensitivity increased for auditory signals on the second test day ($p < 0.01$); no significant change was noted for the visual signals. A significant increase in β was also noted for the auditory signals ($p < 0.01$) but not for the visual signals.

Order effects were also analyzed for each test day. They were not significant—*i.e.* d's and βs did not differ significantly when auditory or visual signals came first. This finding is contrary to that of Experiment I.

TABLE III

MEDIAN β AND d' FOR THE AUDITORY AND VISUAL TASKS IN EXPERIMENT II

Task	β			d'		
	Total Group	A-V*	V-A*	Total Group	A-V	V-A
Session 1						
auditory	2.62	2.73	2.51	1.50	1.43	1.56
visual	2.55	1.84	2.98	1.73	1.99	1.66
Session 2						
auditory	4.86	4.61	9.24	2.33	2.22	2.44
visual	2.93	2.49	4.28	1.92	1.98	1.86

* A-V and V-A refer to the order in which the two tasks were performed.

TABLE IV

COEFFICIENTS OF CORRELATION BETWEEN VISUAL AND AUDITORY TASKS (CM) AND BETWEEN SCORES ON SUCCESSIVE DAYS (TRT) IN EXPERIMENT II

(Correlations were computed separately for Sessions 1 and 2, and for auditory and visual tasks, A and V; r_s and r_p are rank-order and product-moment coefficients, respectively. β was transformed to its common logarithm before r_p was calculated.)

	CM				TRT			
	d'		β		d'		β	
	1	2	1	2	A	V	A	V
r_s	.71*	.59	.49	.35	.69	.73	.63	.64
r_p	.78	.60	.78	.32	.77	.73	.48	.73

* Coefficients >0.53 are significant at the 0.05 level; >0.66 at the 0.01 level.

Coefficients of correlation across modalities and across days are given in Table IV. As in Experiment I, both Spearman and Pearson coefficients were calculated with the Pearson coefficients for β computed after a logarithmic transformation. On the first test-day there was a moderately large correlation between sensitivity for visual signals and for auditory signals, and between the criteria used in each test. These results are similar to those of Experiment I. On the second test day, the correlation for d'

was again high, but that for β was appreciably lower and was not significantly greater than zero.

Test-retest correlations (Day 1 vs. Day 2) were generally moderately high, though the Pearson test-retest correlation for β (0.48), was not statistically significant.

Most of the cross-modality correlations were of an order of magnitude comparable to the test-retest correlations. This would suggest that the coefficients are appreciably lowered by the inherent unreliability of the psychophysical procedures. This should be qualified, however. It is probable that had test-retest correlations been obtained in a period of time as short as that over which the cross-modality correlations were obtained, they would have been somewhat higher.

General Considerations

The results obtained in the two experiments support the position that there are common factors underlying the detection of visual and auditory signals, both a factor influencing sensitivity as well as a common response bias in reacting to the two kinds of signals.

The fact that d' increased significantly on the second day of Experiment II for the auditory task and not for the visual task, and the fact that there was a significant correlation for β on the first day and not on the second indicates that some different processes were involved in the two tasks. It is possible that the difference reflects something in the quality of the pulses employed rather than an effect of modality itself. For a noise of the kind employed, transient fluctuations of intensity and frequency are apparent even within a 0.5 sec. pulse. Presumably *O* somehow averages these in arriving at a judgment. For a light no such fluctuations are noted. The substitution of tonal signals for the white noise might provide a check of this notion. In any event, it should be noted that Swets and Sewall also reported a small, significant increase in d' in early stages of practice for a similar auditory task.[4]

It was noted above that the effect of order found in Experiment I did not appear in Experiment II. It might be concluded that the result in Experiment I was of no significance. This may be premature. The order may have been effective in the first experiment because of the discrepancy in the detectability of the two types of signals. One might argue that since the visual increments were so much more detectable, *O* could afford to be more conservative (as observed) and still detect a considerable number of signals. This set might then transfer to the detection of the less evident auditory signals after it had been once established. In the second experiment, the signals in the two modalities were of roughly comparable detectability, and so no such difference in set would be established. Experiments in which signals are manipulated in difficulty would cast more light on the subject.

Another question should be raised: Is it necessary to analyze these data in terms of detection-theory? What is the relationship of correct to false responses on these auditory and visual tasks?

Table V gives the Spearman and Pearson cross-modality correlations for the number of correct responses and of false responses in both experiments. Note that there is an appreciable and significant correlation for false responses in almost every case but that a significant cross-modality correlation for correct responses was obtained only in Session 2 of the Experiment II.

Possibly, the common factor of sensitivity to increments in intensity rests primarily on the ability to identify correctly the non-signal pulses; possibly, too, it is this ability which changes with increasing experience. A more clear-cut identification of the underlying variables must await additional experiments.

SUMMARY

Two experiments were performed in which *O*s were to detect which pulses of noise or pulses of light were slightly more intense than others in a train. It was found in the first experiment that both sensitivity (*d′*) and degree of conservatism in responding (β) were correlated for the two tasks. The auditory task was more difficult than the visual under the chosen conditions. Higher βs were noted for *O*s first performing on the visual task. In the second experiment, the tasks were more closely equated in difficulty, and the *O*s performed both the auditory and the visual task twice. The effect of order previously noted for β was not obtained. β and *d′* were again significantly correlated for the first session; for the second session significant correlations were obtained for *d′* but not β. The results suggest that not only are there response biases common to detection in different modalities but also common factors affecting sensitivity.

TABLE V

CROSS-MODALITY CORRELATIONS OF DETECTION (D) AND FALSE RESPONSES (FR)

(Coefficients of correlation between auditory and visual tasks using the number of correct responses (D) and of false responses (FR) in Experiments I and II. For levels of significance, see Tables II and IV.)

	Experiment I		Experiment II			
			Day 1		Day 2	
	D	FR	D	FR	D	FR
r_s	.06	.71	.27	.72	.66	.43
r_p	.11	.57	.21	.79	.62	.67

TABLE VI

MEDIAN NUMBERS OF CORRECT RESPONSES (D) AND FALSE RESPONSES (FR)

	Experiment I		Experiment II			
			Day 1		Day 2	
	D	FR	D	FR	D	FR
A	21.5	9	20	16	23.5	7.5
V	30.7	7.5	25	15	25	11.5

ABILITY AND THE USE OF OPTIMAL STRATEGY IN DECISIONS

By ESTELLE SINGER, Tufts University

The two-choice task commonly used in studies of decision-making requires the *S* to predict which one of two alternatives will occur next on each of a long series of trials. Generally, the series is ordered at random, except for specification of the total frequencies of each alternative. The greatest objective success can be achieved on this task by the simple strategy of choosing the more frequent alternative on every trial. However, this strategy is rarely used. Responses most often approximate 'probability matching,' the choice of each alternative in proportion to its probability of occurrence.[1]

Such less-than-optimal behavior is surprising when it is assumed that the *S*s are readily able to formulate the rational solution. It is reasonable, however, to suppose that the solution itself is difficult to achieve. If so, use of the optimal strategy may be associated with unusual quality of performance in other situations involving similar abilities.

The present study investigated this expectation by factor-analyzing scores on performance and on estimation in the two-choice task, scores on other types of tasks also using sequentially-presented information, and scores on more comprehensive intellectual tests. The specific hypotheses were: (1) The various scores which assess ability to utilize sequentially-presented information would appear together on the same factor(s); and (2) The measures of adherence to the optimal strategy on the two-choice task would have been positive loadings on that factor(s).

Method: (1) Subjects. Thirteen men and fourteen women undergraduates at Tufts University participated in the study. They were paid at the rate of $1 an hour.

(2) Tests. The expectation that use of the optimal strategy on the two-choice task would be associated with the quality of performance in other situations involving similar activities was investigated, as mentioned above, by a factor analysis of scores on several tasks also having information sequentially presented, scores on more comprehensive intellectual tests, and scores on the central two-choice task. Accordingly, one group of tests measured ability to utilize sequentially-presented informa-

---

* Revised manuscript received June 29, 1966. Supported in part by Grant G-10947 from the National Science Foundation, and in part by Contract AF19(628)2450 with the Decision Sciences Laboratory, United States Air Force. Computer time was provided by the M.I.T. Computation Center.

[1] H. W. Hake, and Ray Hyman, Perception of the statistical structure of random binary symbols, *J. exp. Psychol.*, 45, 1953, 64-74.

tion. An attempt was made here to require different uses of the information, such as, in rote memory, concept formation, and immediate reproduction, and at the same time to keep the method of presentation uniform. The basic test, the two-choice sequence, attempted to measure conformity with the optimal strategy. In addition, it included other information on *S*'s estimate of probability and his success in guessing, which serves to clarify the central measure. Finally, standardized tests of mathematical and verbal aptitude were used to measure more general intellectual ability. Sex was also included as a variable.

The tests, described below, were administered in one group-session lasting about two hours. They were given in the order described, with a 10-min. break between the *two-choice* and *memory-span* tests.

(a) Rote learning sequence. This task, like the other prediction tasks to be described below, required *S* to predict the next figure in a series. The series could be mastered best by memorization, since it consisted of a short sequence of figures which was repeated 10 times. The procedure was as follows: On each trial *S* selected 1 of 4 alternative figures (2 geometric shapes in 2 colors) as his prediction, recorded it on an answer pad, turned the page, and was given the correct answer. A trial lasted approximately 12 sec., with 7 sec. for the response. Three series were given, each consisting of a unit sequence of figures which was repeated 10 times. The lengths of the unit sequences in the three series were 6, 5, and 4 respectively. In addition, a recognition-test was given after each series which contained several questions of the form, "The first five predictions given by another person are shown below. How many of his predictions do you estimate were correct?" Only five predictions were asked for in order to avoid suggesting to the *S* the repetitiveness of the series. The recognition-test was intended to provide an estimate of learning which was at least partially independent of whatever strategy had been used during the prediction-trials. The two scores that remitted were (1) the total number of correct predictions during the task; and (2) the correct identifications of the items that belonged and did not belong on the recognition-test.[2]

(*b*) *Two-choice sequence.* This task followed the same procedure as the *sequence* in *rote learning,* except that there were only two alternatives from which to choose. The series of 80 trials was constructed randomly, subject to the constraints that the more frequent alternative appear 67% of the time and that the less frequent alternative was never repeated on the next trial. The latter characteristic was included to permit a unique, or nonprobabilistic, solution in a small portion of the trials in order to assess the sensitivity of performance when the optimal strategy was not used. The appropriateness of the optimal strategy itself would not be changed. Estimates were obtained, after the series, of what *S* believed the percentages of occurrence of each alternative to be and of what were the frequencies of each of the four possible sequences of two figures.

Six scores were obtained. The first four concerned the more frequent figure, and consisted of the *direction* and *amount* of deviation of the *individual predictions* and of the later *estimates* from 67%. The last two scores concerned the less frequent figure, and consisted of the frequency of repeated choices and the later estimate of the degree of repetition of that figure.

Specifically, the first, trichotomous score, (3), was based on whether the more frequent alternative was chosen more often, equally often, or less often than 67%.[3] It was designed to assess conformity with optimal strategy, in contrast to 'probability matching' or other more balanced strategies. The second score, (4), was the absolute amount of deviation of the more frequent figure choices from 67%.[3] This score was included for comparison with the comparable score from estimation. The third, trichotomous score, (5), was based on whether the estimate of occurrences of the more frequent figure was over, was accurate, or was under. As with the first score, an overestimate was labelled positive. The fourth score, (6), was the absolute error in the estimation of the percentage of occurrences of the more frequent figure. The fifth, (7), and sixth, (8), scores were the number of repetitions of the least frequent figure and the estimated percentage of repetitions of the least frequent figure. They both reflect error, since the series never repeated this figure.[4]

---

[2] The numbers in parenthesis indicate the scores as listed in Tables I, II, and III.

(c) Memory-span test. This task, modelled after the Digit Span Test, visually presented short sequences of figures at the rate of one figure per sec.[5] *S* was to reproduce the series on his answer sheet. The eight series ranged in length from 4 to 7 figures, and were composed of the same type of colored geometric figures as were used in the other tasks. The score for this task was the total number of figures reproduced in the correct position, (9).

(d) Concept-formation sequence. The procedure for this task was the same as for the other two prediction tasks, except that there were nine possible choices. The 72 trial series was composed of short, repeating sequences of colors (*i.e.* red, green, red, blue) and of shapes (*i.e.* circle, square, and triangle). Considering color and shape together, the unit sequence was of length 12, much greater than the rote memory-span of any of the *S*s. A perfect score could be obtained by examining the two dimensions separately, and recognizing that each repeated a short sequence. The score used was the number of correct predictions of both attributes (10).

(3) *Other measures.* College Entrance Examination Board Verbal (11) and Mathematics (12) scores, and sex (men were scored as positive) (13) were obtained from university records and kept confidential.

Results and discussion. Pearson product-moment correlations were calculated and factor-analyzed by the Jacobi iterative method.[6] Two significant factors were indicated which were then rotated to an orthogonal solution.[7] Table I presents the intercorrelations; Table II, the unrotated factor loadings and the lower bounds of the estimates of the reliability coefficients; and Table III, the rotated factor loadings.

---

[3] Only the latter half of the trials were used in order to avoid initial response-preferences.

[4] Scores (7) and (3) may be experimentally independent in that a given percentage of responses predicting the less frequent figure can be obtained by repeating or not repeating when the percentage is less than 50, as in this case. The two measures may be related, however, in that there is more likelihood of repetition when there are more such responses. It was decided to include both measures when it was found that the range of variation of the total percentage across *S*s was small ($\sigma = 3.04$) thus minimizing the effect of relative frequency on the second, sequential effect.

[5] David Wechsler, *Wechsler Adult Intelligence Scale, Manual,* 1955, 176-177.

[6] H. H. Harman, *Modern Factor Analysis,* 1960. The Pearson product-moment correation was used on the trichotomous scores, (3) and (5). Aternatives, such as the biserial coefficient, were not calculated because of the characteristics of computer programs available to us at the time of analysis. It was judged that the main effect of using a Pearson *r* would be to attenuate the magnitude of the obtained *r*s, and therefore their contributions to the communality. Fortunately, this effect does not appear to have altered the essential factor structure, *e.g.* Factor 1 is defined by a trichotomous score, (3), and several continuous scores, (4), (7), and (12).

The first hypothesis was that the various scores which assess ability to utilize sequentially-presented information would appear together on the same factor. Accordingly, the scores that should have had positive loadings on such a factor were (1), (2), (9), and (10), and those that should have had negative loadings were (6), (7), and (8). Factor 2 confirmed

TABLE I

INTERCORRELATIONS AMONG THIRTEEN MEASURES DESCRIBED IN TEXT

($N=27$)

Scores

Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	—											
(2)	.45	—										
(3)	.20	-.24	—									
(4)	.04	.22	-.54	—								
(5)	.04	-.29	.25	-.23	—							
(6)	-.16	-.34	.16	-.10	-.06	—						
(7)	-.22	.06	-.41	.59	-.23	-.06	—					
(8)	-.23	-.17	-.27	.06	-.15	-.01	.35	—				
(9)	.19	.52	.06	-.20	.09	-.42	.02	-.13	—			
(10)	.05	.40	.23	-.04	-.06	-.41	-.36	-.38	-.05	—		
(11)	.19	.24	-.27	.42	.06	-.39	.49	-.24	.19	.10	—	
(12)	.31	.24	-.33	-.39	-.32	-.08	.36	.11	.31	-.15	.21	—
(13)	.08	.18	-.04	.27	-.37	.09	-.01	-.02	.08	-.12	-.03	.60

TABLE II

UNROTATED FACTOR LOADINGS AND LOWER BOUND OF ESTIMATE OF RELIABILITY COEFFICIENTS

Score	1	2	h^2	Lower bound of reliability
(1)	.16	.67	.47	.45
(2)	-.40	.72	.68	.52
(3)	.69	.14	.50	.54
(4)	-.76	-.02	.58	.59
(5)	.46	.05	.21	.37
(6)	.25	-.60	.42	.42
(7)	-.73	-.24	.59	.59
(8)	-.25	-.50	.31	.38
(9)	-.17	.57	.35	.52
(10)	.20	.75	.60	.41
(11)	-.52	.35	.39	.49
(12)	-.72	.02	.52	.60
(13)	-.43	-.08	.19	.60

---

[7] H. F. Kaiser, The varimax criterion for analytic rotation in factor analysis *Psychometrika* 23, 1958, 187-200.

the hypothesis in that all but one of the above scores showed substantial loadings and these were in the proper direction. Only one additional score (11) showed a loading on this factor, and this was the lowest considered in the table. The more general intellectual measures of mathematical aptitude (12) and verbal aptitude (11) did not contribute substantially to Factor 2. Factor 2 may therefore describe a specific ability to assess and utilize sequentially-presented information.[8]

The loading, albeit low, of the verbal aptitude measure on Factor 2 may indicate that properties of sequences in language are similar to those in the tasks used in this study. The correlations suggest that accuracy of estimation in probabilistic sequences is the most relevant common property. If this suggestion were confirmed by further research, it would add support to the information theory approach to the investigation of language.[9]

TABLE III

ROTATED FACTOR LOADINGS

Score	Measure	Loadings	
		Factor 1	Factor 2
(1)	Correct predictions of rote learning sequence	—	.65*
(2)	Correct identifications of rote learning sequence	—	.76
(3)	More frequent figure predictions greater than 0.67 on two-choice sequence	−.70	—
(4)	Absolute deviation of more frequent figure predictions from 0.67 on two-choice sequences	.75	—
(5)	More frequent figure estimate greater than 0.67 on two-choice sequence	−.46	—
(6)	Absolute deviation of more frequent figure estimate from 0.67 on two-choice sequence	—	−.62
(7)	Repetitions of least frequent figure on two-choice sequence	.76	—
(8)	Estimated repetition of least frequent figure on two-choice sequence	—	−.47
(9)	Correct recall on memory span test	—	.58
(10)	Correct predictions of concept formation sequence	—	.72
(11)	CEEB Verbal Score	.48	.41
(12)	CEEB Mathematics Score	.71	—
(13)	Sex (men)	.44	—

* The loadings included are greater than |0.33|, the magnitude of a correlation just significant at the 0.05 level.

The second hypothesis was that the measures of adherence to the optimal strategy on the two-choice task would have positive loadings on the factor just discussed. The primary measure of optimal strategy, (3), showed a loading instead on an independent factor, Factor 1. That factor, it appears, is defined by measures on the two-choice task, of general intellectual ability, and by sex, but not by measures on the other tasks involving sequentially-presented information.

---

[8] Since the study was primarily designed to investigate the two-choice situation, it quite evidently did not provide a wide enough variety of tests to identify this factor fully.

[9] Fred Attneave, *Applications of Information Theory to Psychology,* 1959.

The loadings contributed to Factor 1 by the two-choice task tended to be measures of the less optimal performance. Thus, there were fewer choices of the more frequent figure, (3) and (4), and more repetition of the less frequent figure, (7), which had never repeated in the series. At the same time, however, the loadings contributed by the more general intellectual tests were measures indicative of superior intellectual performance.

The quality of performance on the two-choice task appears, therefore, to be associated *negatively* with general intellectual aptitude, as measured by our tests. It appears to be *independent* of the more specific ability revealed by Factor 2, which may involve use of sequentially-presented information. The latter conclusion must be qualified to the extent that error of estimation has variance in common with the tasks involving sequential information.

Explanations which have been offered of why the optimal strategy does not usually appear in two-choice tasks can be examined as follows. (*a*) *S* does not want to settle for less than a perfect score, and therefore prefers to attempt to predict even the infrequent event sensitively.[10] If this were the case, measures (7) and (8), reflecting error in prediction of this event, would have had loadings in the same direction as measure (3), which reflects use of the optimal strategy. (*b*) The *S*s attempt to find a unique solution rather than working on the basis of a probabilistic environment.[11] If this were the case, measures (7) and (8), which involve a unique solution, would be expected to have shown loadings on Factor 2 which contains other predictive tasks permitting of a unique solution. (*c*) The *S*s attempt to take account of as much information as possible, and therefore attend to more subtle aspects of the series than the over-all frequencies.[12] If this were the case, the accuracy of judgments about the composition of the series made after the prediction trials might (1) have shown loadings on Factor 2 which also involves close attention to the composition of the series, and (2) have shown loadings on Factor 1 in the opposite direction from the measure of use of the optimal strategy (*i.e.* *S*s having used the optimal strategy would not have made accurate judgments, and *vice versa.*)

Positive loadings on Factor 1 include both less frequent choice of the more frequent alternative and underestimation of the proportion of occurrence of that figure. The apparent correspondence between the direction of the choices and the estimates suggests that Factor 1 describes probability-matching based upon estimates of probability having varying degrees of bias. The results, therefore, imply that the *S*s respond directly to each alternative in proportion to the perceived frequency of success. The more 'rational' *S*s may be those who overestimate the relative frequency of the predominant figure.

---

[10] J. S. Bruner, J. J. Goodnow, and G. A. Austin, *A Study of Thinking,* 1956, 194.

[11] M. W. Flood, Environmental non-stationarity in a sequential decision-making experiment, in R. M. Thrall, C. H. Coombs, and R. L. Davis (eds.), *Decision Processes,* 1954, 287-299.

[12] E. S. Singer and T. B. Roby, Measurement of sequential behavior by successive differences, this JOURNAL, 80, 1967, 276-281.

It is paradoxical that the more 'rational' *S*s were less able on the general intellectual aptitude-tests. It may be that the less intelligent *S*s were willing to make more extreme judgments about the magnitude of the proportions. This aspect of the results suggests that future research on decision making needs to consider style of response.

SUMMARY

Optimal behavior on a two-choice prediction-task was related to performance on other tasks involving sequentially presented information and on tests of general intellectual aptitude. As predicted, one factor found in a factor analysis of the results included the measures of how well sequential information was used. Contrary to the prediction that optimal performance on the two-choice task would show loadings on that factor, the quality of performance on the two-choice task showed loadings on an independent factor; the scores for intellectual aptitude also showed loadings on that factor but in the opposite direction. It was tentatively concluded that the two-choice task elicits a direct form of probability-matching, based on subjective estimates of the probabilities.

THE EFFECT OF FALSE PERSPECTIVE CUES ON DISTANCE- AND SIZE-JUDGMENTS: AN EXAMINATION OF THE INVARIANCE HYPOTHESIS

By W. W. BLESSING, University of New South Wales,
A. A. LANDAUER, University of Western Australia,
and MAX COLTHEART, Monash University

An object of physical size S located at physical distance D from the observer O can be said to have a perceived absolute size S' and a perceived absolute distance D'. According to the size-distance invariance hypothesis (*SDIH*), $S' = k \Theta D'$, where k is some constant and Θ is the visual angle of the object.[1]

Recently, Epstein, Park, and Casey reviewed the literature dealing with this hypothesis and concluded that the available evidence did not enable a firm decision to be made concerning its tenability.[2] They pointed out that investigating the conformity of group data to this hypothesis is of questionable value in cases where individual differences are such that there are wide discrepancies between individual data and predictions from the *SDIH*. They also recommended that in studies of the *SDIH* paired judgments of size and distance ought to be obtained, since the relationship of D to D' may not be such that D can be substituted for D' in the above equation.

Since then, Hartman, and Gogel, Wist, and Harker, have reported experiments in which some support for the *SDIH* was found.[3] Gogel, Wist, and Harker studied the effect on S' of varying D' by using an optical device to increase or decrease the effective interpupillary distance of O.[4] They found that S' and D' were positively correlated (as required by the *SDIH*) but that the ratio S'/D', with Θ constant, was *not* invariant over a range of values of D'.

In the present study, false perspective is used for producing variations in D'. It was intended that in one condition $D' < D$, and in another condition $D' > D$, with D constant in both conditions. Under these conditions, the size-distance invariance hypothesis may be tested in various ways.

---

* Received for publication July 25, 1966. From the University of Sydney, Sydney, Australia.

[1] F. P. Kilpatrick and W. H. Ittelson, The size-distance invariance hypothesis, *Psychol. Rev.*, 60, 1953, 223-231.

[2] William Epstein, John Park and Albert Casey, The current status of the size-distance hypothesis, *Psychol. Bull.*, 58, 1961, 491-514.

[3] A. M. Hartman, Effect of reduction on the relationship between apparent size and distance, this JOURNAL, 77, 1964, 353-366.

[4] W. C. Gogel, E. R. Wist, and G. S. Harker, A test of the invariance of the ratio of perceived size to perceived distance, this JOURNAL, 76, 1963, 537-553.

METHOD

Apparatus. Variation of apparent distance was achieved by means of two frustra of right square pyramid-shaped viewing tunnels (see Fig. 1). *O*'s end (viewing end) of the tunnel was a 25 × 25 cm. square in the case of Tunnel A and a 75 × 75 cm. square in the case of Tunnel B. For both tunnels the other end (stimulus-end) was a 50 × 50 cm. square and the perpendicular distance between viewing- and stimulus-ends was 450 cm. The tunnels were made of hard-board and their interior was painted flat black. Nine yellow lines were painted latitudinally around the interior of each tunnel. The position and width of these lines is shown in Table I. The viewing end of each tunnel was covered by a masking board and the tunnels themselves were screened from *O*'s view by means of black drapes. The only view of the tunnel which *O* could obtain was through a viewing aperture (11-mm. diameter), placed in the center of the viewing end, at the eye-level of a seated *O*.

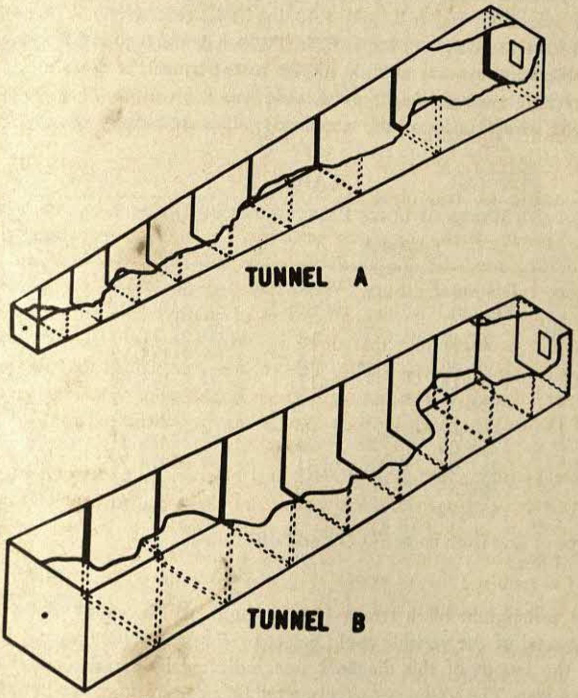


FIG. 1. THE VIEWING TUNNELS

In the case of Tunnel A, the display presented to *O* was consistent with the linear perspective of a rectangular tunnel (closed at both ends) 225 cm. long with latitudinal lines of equal width spaced at equal intervals. The display presented by Tunnel B was consistent with the linear perspective of a rectangular tunnel 675 cm.

long, again with latitudinal lines of equal width spaced at equal intervals. An identical rectangular piece of white cardboard, 16.4-cm. high and 11.1-cm. wide (20-cm. diagonal), was attached to the center of the stimulus-end of each tunnel and served as standard. Its diagonal subtended a visual angle of 0.0444 radians (2°32′).

The tunnels were illuminated by two 100-w. incandescent lamps placed in parabolic reflectors situated at the top right and lower left corners of the viewing end of each tunnel. These lights could not be seen by *O*. The reflectance of the first line in Tunnels A and B (as measured by a Salford Electrical Instruments Ltd. photometer) was 1.48 log ft-L. and 1.40 log ft-L. respectively. The reflectance of the standard stimulus was 0.42 log ft-L. in Tunnel A and 0.40 log ft-L. in Tunnel B.

The variable stimulus was at right angles to the tunnel, at *O*'s right. It consisted of a white rectangle whose height-width ratio was fixed at 3.1:2.1 (which was equal to the height/width ratio of the standard). This rectangle was attached to the center of a rectangular black screen (79-cm. high, 48-cm. wide) at the eye-level of *O*. The diagonal of the variable could be set by *E* at any value between 3.75 cm. and 37.5 cm.; the length of this diagonal was indicated by a pointer and scale at the rear of the screen. The screen was so connected by means of stays to a movable base that the variable could be positioned anywhere along a 900-cm. long and 140-cm. wide corridor. A door-way 550-cm. from *O*, drapes and the usual laboratory illumination by fluorescent light gave many distance cues.

TABLE I

POSITION AND WIDTH OF LINES PAINTED ON THE INTERIOR OF THE TUNNELS

	Tunnel A		Tunnel B	
Line	Distance* (in cm.)	Width† (in mm.)	Distance* (in cm.)	Width† (in mm.)
1	26.49	13	71.10	35
2	56.33	15	135.12	32
3	90.13	17	193.08	29
4	129.07	19	245.67	26
5	173.24	22	293.81	24
6	255.66	26	338.11	22
7	287.93	31	378.72	20
8	360.55	36	416.67	19
9	450.72	45	450.81	17

* Distance of line from corner at *O*'s end of tunnel.
† Width of line.

Observers. Thirty *O*s (16 men and 14 women), all undergraduates with uncorrected vision, attending an elementary course in psychology, were randomly assigned to the two experimental conditions, 15 viewing the standard in Tunnel A and 15 viewing it in Tunnel B.

Instructions. In the formalized instructions, emphasis was placed on the necessity that distance- and size-judgments be *objective*, that all judgments be independent of the preceding ones, and that head movement be minimized. The *O*s were then instructed to place their foreheads against the masking board and to look at the stimulus through the viewing aperture, using their left eyes only and taking care to keep *both* eyes open.

Procedure. For both experimental conditions, the variable was first so covered that only the black screen was visible; the screen was then placed in an initial position which was randomized over the course of the experiment. Four judgments of the distance of the standard stimulus were then made by each *O*, *E* moving the screen backward or forward as directed by *O*, till *O* was satisfied that the variable and the standard were equidistant from him. Two judgments in ascending and two in descending order were made, counterbalanced over *O*s. Between successive judgments, *O* was instructed to keep looking at the standard stimulus for the time it took *E* to reposition the variable. After *O* made his distance judgments, *E* placed the screen at the mean distance judged and uncovered the variable. *O* was then instructed to make two ascending and two descending size-judgments (also counterbalanced), *E* increasing or decreasing the size of the variable as directed by *O* until *O* was satisfied that the variable was the same size as the standard. After each judgment, the length of the diagonal of the variable stimulus was recorded by *E*. The *O*s were again reminded to give independent judgments and to look at the standard stimulus between judgments.

TABLE II

MEANS, ACROSS MEASURES, FOR THE INDIVIDUAL *D′* AND *S′* (in cm.)

Tunnel A		Tunnel B	
D′	*S′*	*D′*	*S′*
399.50	16.50	286.50	15.13
323.75	16.80	590.00	22.78
148.00	10.00	487.00	27.98
252.00	10.80	390.25	14.35
335.50	16.08	453.50	22.63
324.75	19.43	687.50	29.25
351.50	23.05	362.00	17.63
300.75	15.60	400.25	16.90
477.00	22.30	598.75	24.83
285.75	15.27	387.50	8.30
484.50	20.08	576.50	30.98
386.50	19.40	429.50	27.30
303.00	9.53	335.00	18.08
278.50	14.96	413.50	17.38
338.00	15.98	495.50	23.45
M = 332.43	M = 16.38	M = 459.42	M = 21.38
S.D. = 80.88	S.D. = 3.98	S.D. = 108.75	S.D. = 5.25

RESULTS

Table II shows the individual *D′* and *S′* values for the two experimental conditions, as well as means and standard deviations of these values. The individual entries on the table were obtained by computing the mean of the four judgments made by each *O*.

(1) Apparent length of tunnels. As stated earlier, the false perspective was intended to make Tunnel A look shorter than it actually was and Tunnel B longer than it actually was. Since the physical length of both

tunnels was 450 cm., the effectiveness of the false perspective cues can be evaluated by testing, for each tunnel separately, the hypothesis

$$H_0 : \mu_{D'} = 450 \text{ cm.},$$

where $\mu_{D'}$ is the mean of the population values of D′ for each tunnel. A test of H_0 for Tunnel A yielded a *t*-value of 3.51 ($p < 0.01$); therefore, $D' < D$.

For Tunnel B, H_0 yielded a *t*-value of 1.02 ($p > 0.05$) and the hypothesis is retained. In the case of Tunnel B,D' and D do not differ significantly. Nevertheless, one can deduce from these results that D' differed significantly for the two tunnels and that consequently the two experimental groups used were exposed to different values of D'. A *t*-test of the difference between the mean apparent lengths of the tunnels would have indicated that these two means are significantly different, as is required, but would not have allowed separate assessments of the effectiveness of the false perspective cues in each tunnel.

(2) Apparent size of stimulus. Since for Tunnel A, D' was less than for tunnel *B*, the *SDIH* predicts that S' will be less for Tunnel A than for Tunnel B; *i.e.* it predicts that hypothesis

$$H_0 : \mu S'_A - \mu S'_B = 0$$

will be rejected.

The value of S'_A was 16.38 cm. and of S'_B was 21.38 cm.; a test of this difference yielded a *t*-value of 2.86 ($p < 0.01$). Consequently, H_0 is rejected, confirming the prediction from the *SDIH*.

(3) Correlations between S′ and D′. The product-moment correlations, between the S' and the D' value for each individual *O* were as follows: for Tunnel A, + 0.77, for Tunnel B, + 0.74, and for the combined data, + 0.79. For each correlation, the hypothesis

$$H_0 : \rho = 0.00$$

was tested. Assuming that a sample of pairs of observations is randomly drawn from a bivariate normal distribution, then, if the sample correlation coefficient is *r*, the quantity

$$r^2/(1 - r^2)\ (n - 2)$$

has an *F*-distribution with 1 and $(n - 2)$ *df*. The *F*-value for each correlation was significant beyond the 0.01 level.

McNemar gives a procedure for evaluating departure from linearity in a correlation situation.[5] This procedure involves calculating *F*-ratios whose magnitude depends on the size of the difference between the correlation ratio *eta,* and the product-moment coefficient, since the degree of nonlinearity is reflected by the size of this discrepancy.

---

[5] Quinn McNemar, *Psychological Statistics*, 3rd ed., 1962, 275-278.

The values of *eta* for the combined data were + 0.86 and + 0.82 and these values did not lead to significant *F*-ratios. It can therefore be concluded that there is no evidence of a nonlinear component in the relationship between *S′* and *D′*.

Fig. 2 shows the relation between *S′* and *D′*: the line of best fit (least squares) closely approximates the theoretical visual angle function, *i.e.* the function specified by the *SDIH* with *k* equal to unity.

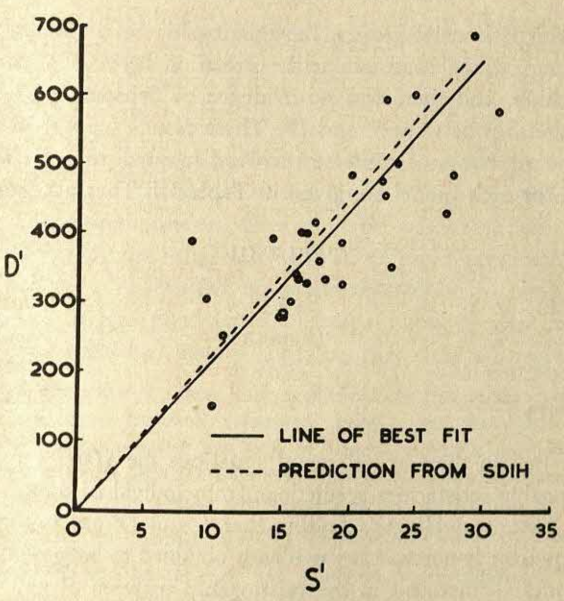


FIG. 2. THE RELATION BETWEEN PERCEIVED ABSOLUTE SIZE (*S′*) AND PERCEIVED ABSOLUTE DISTANCE (*D′*) IN CM.

(4) Values of the ratio S′/D′ for the two groups. The *SDIH* may be formulated thus:

$$S'/D'\Theta = k$$

Gogel, Wist, and Harker found that there was an inverse relationship between group mean values of *k* and *D′*, a result which is inconsistent with the *SDIH*.[6] In the present experiment, the stimulus in each tunnel subtended a visual angle of 0.0444 radians along its diagonal, and therefore the value of *k* for each Θ can be calculated by the formula

---

[6] Gogel, Wist, and Harker, *op. cit.*, 553.

$$k = S'/D' \cdot 1/0.0444$$

For Tunnel A, the mean value of k was 1.13; for Tunnel B, the mean value of k was 1.04. Since $D'_A < D'_B$, the difference between these values was in the direction to be expected from Gogel, Wist, and Harker's results. The values of k, however, were not significantly different. ($t = 1.12$, $p > 0.05$).

DISCUSSION AND CONCLUSION

For both experimental groups, the relationship between S' and D' was a positive one, the product moment correlation between S' and D' was reasonably high, and there was no evidence of departures from linearity in the relationship between S' and D'. These results support the *SDIH*.

The standard errors of estimate involved in predicting S' from D', or D' from S' for each tunnel are given in Table III. They are small enough to indicate that in these experimental situations the *SDIH* can be used to make reasonably satisfactory predictions from individual data.

TABLE III

STANDARD ERROR OF ESTIMATE (in cm.)

	Tunnel A	Tunnel B
S' predicted from D'	2.55	3.54
D' predicted from S'	51.76	73.35

Gogel, Wist, and Harker's finding that k and D' are inversely related was not confirmed; nor was any evidence obtained to suggest the presence of a nonlinear component in the relationship between S' and D' such as they suggested.[7] Since, however, the direction of the difference between the obtained values of k was as predicted by Gogel, Wist, and Harker, it would seem that further research might profitably be directed at studying in greater detail the relationship between D' and k.

SUMMARY

The relationship between perceived absolute size S' and perceived absolute distance D' was examined by using false perspective cues to produce illusory variations in D'. No evidence was found to suggest that S' and D' are not related in the manner specified by the size-distance invariance hypothesis. It was found that this hypothesis can be used to make accurate predictions of S' from D' and vice versa for individual *O*s.

---

[7] Gogel, Wist, and Harker, *op. cit.*, 544.

RELAXATION AND THE PERKY EFFECT: THE INFLUENCE OF BODY POSITION ON JUDGMENTS OF IMAGERY

By SYDNEY J. SEGAL and MICHAEL GLICKSMAN, New York University

Perky's early study has often been interpreted as indicating that *O*s were unable to distinguish introspectively between an image and a weak perceptual experience.[1] Working in Titchener's laboratory, Perky instructed *O*s, some of whom were trained introspectionists, to imagine an object while they were looking at a screen on which there was projected a very weak colored light in the shape of the object. Thus, for example, the *O* was asked to imagine a tomato when the light was round and red. While the stimulus was bright enough so that an *O* could presumably see it if he had expected it, Perky's *O*s, under the instruction to imagine the object, all failed to recognize that an external stimulus was present; and when told that there had been a stimulus, they were all surprised.

The related literature is sparse, but confusions of imagination with veridical perceptions have been frequently observed. Imagery and hallucinations are reportedly facilitated by reducing the level of arousal, although the nature of these experiences varies under different conditions. Silverman, Cohen, Bressler, and Shmavonian reported on imagery during two hours of sensory deprivation, and concluded:

> There may be no such thing as a true hallucination in the sense that no sensory cue is present. Indeed, it may be that all of these experiences originate with some sensation, either internal or external, and subsequently are elaborated.[2]

A curious illustration of this point occurred in their experiment, when *E* inadvertently permitted tiny leaks of light to appear around the door toward the end of the isolation period. The pattern of shadows thus created was described by two *O*s as a veridical perception; but one, a medical student, assumed they were X-rays, while a psychologist believed that a Rorschach card was shown. The *O*s here apparently experienced both an increase of imagery and an impaired ability to make correct perceptual judgments.

Ziskind *et al.* suggested that the different types of imagery and other unusual

---

* Received for publication April 18, 1966. This study was supported by NIMH grant No. MH 06733. From the Research Center of Mental Health. Our thanks are due Drs. G. S. Klein, Leo Goldberger, and Morris Eagle for their assistance.

[1] C. W. Perky, An experimental study of imagination, this JOURNAL, 21, 1910, 422-452. For examples of the use of Perky's observations, see Robert Leeper, Cognitive processes, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 732, and C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 640-641, 654.

[2] A. J. Silverman, S. I. Cohen, Bernard Bressler, and B. M. Shmavonian, Hallucinations in sensory deprivation, in L. J. West (ed.), *Hallucinations*, 1962, 131.

behavior reported during sensory isolation experiments may derive from a single source, and therefore can be considered a syndrome.[3] The hypnoid syndrome is characterized by a diminution of logic, attention, and inhibitory control; and the weakening of these factors leads to poor reality testing, hallucinatory, dream-like, and simple optic imagery, confusion, anxiety, and involuntary movements. Ziskind *et al.* believe that the important cause of the syndrome is not specific conditions of sensory isolation or deprivation, but reduced awareness itself.

These and other studies suggested that a condition of lowered arousal would contribute to blurring accurate judgment in the Perky situation, both by affecting the images experienced by the *O*s and by diminishing their ability to discriminate imagery from veridical perception.[4] The Perky technique provides a low-level stimulus which produces an ambiguous experience capable of being judged either as an image or a perception. Reducing stimulus-intensity makes it less likely that *O* will attribute the experience to the presence of an actual external stimulus. At the same time, it appears that *O* must focus on his imagery if the experience is to be reported as an image. In the few instances where *O* stated that he "never had an image," the experience was always accurately attributed to an external source. Moreover, more 'relaxed' *O*s, that is, the *O*s who 'reacted' to a placebo described as a relaxant, experienced more images than either the *O*s who failed to react to a placebo or the control *O*s.[5]

In the present experiment, lowered arousal was encouraged by placing *O* in a supine, as compared to a sitting and standing, position. To monitor the effectiveness of the changes in posture, readings of body temperature were taken, as these seem to yield a good overall measure of relaxation.[6] It was predicted that the more relaxed body posture would be conducive to lowered arousal and greater 'confusion' in Ziskind's sense. The *O*s in this state would then have more imagery and more vivid imagery, which might increase the likelihood of judgment that they were experiencing imagery despite the presence of the stimulus. It was also predicted that when relaxed, *O* would be less sensitive to stimulation and that this would further diminish the tendency to recognize the stimulus as external. Sitting and standing postures would be accompanied by more adequate reality testing, and there would be greater accuracy in judging the external locus of the stimulus.

---

[3] Eugene Ziskind, Harold Jones, William Filante, and Jack Goldberg, Observations on mental symptoms in eye-patched patients: Hypnagogic symptoms in sensory deprivation, *Amer. J. Psychiat.,* 116, 1960, 893-900; Eugene Ziskind, R. W. Graham, Leonard Kunibou, and Robert Ainsworth, The hypnoid syndrome in sensory deprivation, in Joseph Wortis (ed.), *Recent Advances in Biological Psychiatry,* 1963, 331-345; Eugene Ziskind and Theodore Ausberg. Hallucinations in sensory deprivation—method or madness, *Science,* 137, 1962, 992-993.

[4] Ian Oswald, *Sleeping and Waking,* 1962, 66-87.

[5] S. J. Segal and Shifra Nathan, The Perky effect: Incorporation of an external stimulus into an imagery experience under placebo and control conditions, *Percept. and mot. Skills,* 18, 1964, 385-395.

[6] Nathaniel Kleitman, *Sleeping and Wakefulness,* rev. ed., 1963, 58-63, 131-162.

METHOD

Subjects. Twenty-six women, aged 16 to 25, served as *O*s. About half the group were undergraduates at New York University, who were fulfilling a course requirement by participating. The other half were from several different colleges, and volunteered for the experiment.

Materials. Eighteen 35-mm. Kodachrome slides of simple objects were prepared by reproducing a commercial photograph of the object against a black background. The criteria of selection were that the object have a distinctive but simple shape and color. The slides were then divided into three groups, equated approximately for clarity and brightness of print and general content area. The slides in Group I depicted a red truck, an onion, a pack of Kent cigarettes, a pair of rimless eyeglasses, a white cup of coffee, a roast turkey; those in Group II were a blue car, a pear, a cigarette lighter, a yellow flower, a glass of orange juice, a plate of roast beef; and those in Group III were an elephant, an apple, a silver pen and pencil set, a key, a pilsener glass of beer, and a glass of milk.

Apparatus and procedure. O's axillary temperature was monitored at the start of the experiment, and at 1-min. intervals throughout the procedure, using a direct-reading tele-thermometer, with a thermistor probe held in position throughout the experiment.[7]

O was placed in a translucent plexiglas hood which measured 38 in. in length, 22 in. in width, and extended from just above her knees to 6 in. above her head. This provided a uniform *Ganzfeld,* and virutally eliminated random visual stimulation. The back of the hood was constructed of flexible vinyl, so *O* could lie on it in the supine position. For the sitting and standing positions, the hood could be raised and lowered on a pulley from the ceiling. For the supine position, a mattress was placed on the floor, with the head end raised slightly by a pillow, and the hood rested on the mattress. The order of the three positions was varied among the *O*s.

The small bare experimental room was illuminated, at about 8 f.c. fluorescent lighting throughout the experiment. Next to this room was a small projection-room, and in the wall common to both rooms was a window covered by a sliding wooden panel. Two 300-w. slide projectors were placed in the projection-room: one was aimed at the hood when it hung from the pulley, the other was on a high shelf and was aimed down toward the mattress to project on the hood when *O* was supine. This reduced distortion of the slide in this position. A voltage transformer (variac) was connected to the projector in use, in order systematically to vary the brightness of the slides. The slides were manually controlled, and the variac was operated by an assistant in the projection-room. Sounds from the projection-room were not audible in the experimental room, and the panel in front of the window was always closed except when *O* was under the hood.

O remained in each new position for 7 min. before any reports about images were requested. During this time, *O* had an opportunity to adapt to the hood and the temperature readings became stable in the new position. Then *O* was asked to describe her image of a specific object, *e.g.,* a car. As soon as she indicated that she 'had' the image, the slide representing that object was projected on the hood, the picture was increased in brightness to a supraliminal level;[8] it was reduced again, and extinguished. The total exposure time of each slide was under 2 sec. A total of six slides, one group, was presented in each body position, and each of the three groups of slides was used equally often in each position.

---

[7] Manufactured by Yellow Springs Instrument Co., Yellow Springs, Ohio.

During the experiment, *E* recorded *O*'s reports of imagery and all other remarks. He also recorded his observations of her gestures, posture, and reactions, as a basis for assessing her relative relaxation or tension during the experiment.

After 18 images were obtained, 6 in each position, *O* was removed from the hood and told that "some of the images which I asked you to describe were accompanied by actual pictures which were projected onto the hood." *E* then read over *O*'s description of each image, and asked her to rate it on a five-point scale: (1) I think there was no picture projected; this was my own image; (2) I was somewhat unsure at the time, but now, looking back, I think there was no picture; (3) I wasn't sure at the time, and now, looking back, I'm still not sure; (4) I wasn't sure at the time, but now, looking back, I think there was a picture; and (5) Something was projected, and I knew it all along.

TABLE I

MEAN SCORES FOR IMAGERY, BODY TEMPERATURE, AND ALERTNESS SELF-RATINGS IN THREE BODY POSITIONS

	Standing	Sitting	Supine	*N*	*p**
Imagery score	3.21	2.46	1.43	26	.001
Imagery score, relaxed *O*s	4.24	2.53	1.22	13	.001
Imagery score, tense *O*s	2.18	2.44	1.64	13	n.s.
Body temperature	98.53	98.34	98.13	26	.001
Body temperature, relaxed *O*s	98.57	98.28	98.03	13	.001
Body temperature, tense *O*s	98.56	98.41	98.23	13	n.s.
Self-ratings, alertness	3.44	3.88	4.76	25†	.025
Self-ratings, relaxed *O*s	3.42	4.08	5.25	12†	.025
Self-ratings, tense *O*s	3.46	3.69	4.31	13	n.s.

* The value in each case is based on an analysis of variance.

‡ One *O* in this group used the self-rating scale in a non-scorable manner, and so was excluded from the analysis.

The *O*s were discouraged from using Category (3). In fact, they were most likely to use Category (1) or (5). For each body position, each *O* judged six images and her score for each position is the mean rating. Finally, each *O* rated herself on a seven-point scale of alertness from 7, extremely alert, to 1, extremely relaxed, making a separate judgment for each body position.[9]

RESULTS AND DISCUSSION

The order of the three different body-positions produced no systematic effect, nor did the particular group of slides that was used. Therefore, the data will not be reported separately for these variables.

The predicted main effects appeared, as can be observed in Table I.

---

[8] The supraliminal level was bright enough that three different *O*s could readily see the stimulus, when they were alert and anticipating a discrimination.

[9] Harry Fiss, State of consciousness and the subliminal effect. Unpublished Doctoral dissertation, New York University, 1961, 35-36.

Significantly more reports of imagery occurred in a supine position (low numbers), while the fewest occurred when the *O*s were standing. Body temperature was lowest in the supine position, indicating least arousal, intermediate in the sitting, and highest in the standing position. Self-ratings also pointed to experiences of greater relaxation in the supine position. The correlation between changes in body temperature and judgments that *O* had seen only his own images was positive but non-significant (0.24).

Dividing the *O*s into relaxed and tense subgroups permits a closer examination of the effects. Based on his record of *O*'s impressions of her reactions during the experiment—her physical restlessness and comments ranging from matter-of-fact acceptance of the experimental conditions to a tense and dismayed questioning of every request—*E* divided the 26 *O*s into the 13 more relaxed and the 13 more tense. Perhaps 15 to 20 of the *O*s could easily be assigned to one category of the other; for the rest, the decision was less certain. Some confirmation of *E*'s impression may be adduced from the fact that the mean body temperature was lower (more relaxed) for the 'relaxed' subgroup, and the mean self-ratings also indicate greater relaxation for these *O*s (see Table II).

TABLE II

MEAN SCORES FOR THE TENSE AND RELAXED SUBGROUPS ON IMAGERY, TEMPERATURE, AND SELF-RATINGS

	Tense *O*s	Relaxed *O*s
Imagery scores	2.1	2.7*
Body temperature	98.40	98.30*
Alertness self-ratings	3.82	4.25*

*A *t*-test shows these scores to be significant, $p<0.001$

Dividing the *O*s into these subgroups indicated that the main effect was almost entirely carried by the relaxed *S*s. The imagery scores ranged from 4.24 in the standing position to 1.22 in the supine position, on a scale with 5 and 1 the extreme values. The temperature and self-ratings, likewise, were more extreme, and the correlation of imagery with body temperature was now 0.78 (significant, $p < 0.01$). For the tense subgroup, however, none of the differences among scores in the three positions were significant, and there was no relationship ($r = 0.01$) between temperature and imagery scores.

The effects of tension are not clear in the present experiment, possibly because the distinction between anxiety, tension, and alertness was not clear. This is an area for further investigation. But the main results accord with the hypothesis: More relaxed body postures encourage experiences of imagery and blur judgment about external conditions, while the more alert body-posture is associated with judgments more in accord with reality. Within *O*s, the tendency to report imagery at the expense of accurate appraisal of reality increased with greater relaxation. This was true for 17 of the 26 *O*s. Between *O*s also, the more relaxed reported imagery more frequently, as is shown in Table II.

It is possible that the postural change altered the threshold value for the stimulus, thereby affecting the report of imagery. When *O* is supine, a brighter stimulus may be required to generate a judgment of a veridical perception. Sensitivity to external events may be reduced, and *O*'s attention shifts to inner experiences; images are experienced as more vivid, and faint stimuli can readily be assimilated to the image without conscious perception. In a sitting or standing position, even while viewing the *Ganzfeld* of the hood, *O* appears to be less attentive to inner experiences; he is directed outward and identifies faint stimuli.

In a drowsy or sleeping state, a stimulus may have an effect on a person's imagery without being recognized as a stimulus. In the Silverman example, the shadows were correctly judged to be external, but the *O*s added their own interpretation. In this view, then, the threshold varies as a function of the momentary state of the organism, for example, the relative tension or relaxation. The threshold, however, is a mean value, with a statistical range; and, between the point where *O* can identify a perception accurately and the point where there is a complete lack of awareness of external stimuli, there exists a range of ambiguous experiences.

Within this range, if *O* fails to recognize a stimulus as external, it may still enter the system as 'subliminal' input, and be assimilated to the current train of thought. Still within the range, but closer to the threshold for accurate perceptions, the stimuli are recognized as perceptions, but not identified with complete accuracy: at this level, the current train of thought or imagery may determine the context in which identification occurs.

The effects that we are reporting are clearly related to the "hypnoid syndrome" described by Ziskind. While the Perky *O*s are not as close to sleep as in a true hypnagogic state, nevertheless, the use of the *Ganzfeld* hood creates a condition of modified sensory restriction. Add the instructions to imagine something, and there will be created a condition of reduced awareness.[10] Placing *O* in a supine position may make her more relaxed and still closer to the hypnagogic state. The simultaneous increase in imagery, increase in their vividness, diminution of reality-testing, and raising of the perceptual threshold in this state, all contribute to the effect. It is difficult to isolate the separate contributions of these effects, as all tend to covary along with level of awareness.

---

[10] On the question of instructions, see Ziskind, Significance of symptoms of sensory deprivation experiments due to methodological procedures, in Joseph Wortis (ed.), *Recent Advances in Biological Psychiatry*, 1963, 111-117.

MONOCULAR AND BINOCULAR SIZE-MATCHING AS A FUNCTION OF DISTANCE AT VARIOUS AGE-LEVELS

By H. W. LEIBOWITZ, S. W. POLLARD and D. DICKSON,
Pennsylvania State University

It has been reported recently that neither temporary nor permanent monocularity influences size-matching in a cue-rich environment.[1] A comparison of these results with those of previous studies, in which both temporary[2] and permanent monocularity[3] resulted in a reduction in size-constancy, suggests that experienced *Os* are able to utilize monocular cues to compensate for the lack of binocular vision when many monocular cues are available, but are not able to do so when the number of such cues is diminished. As a test of this hypothesis, the present study was designed to investigate monocular and binocular size-matching in a cue-rich environment for children of various ages as well as for adults. Since young children demonstrate incomplete size-constancy for distant objects,[4] it might be expected that loss of binocular cues would be particularly critical under those conditions for which size-constancy is incomplete or marginal.

EXPERIMENT I

Method: Observers. The *Os* were 90 elementary-school children, grades kindergarten through 6, and 10 college students. There were 20 *Os* in the kindergarten and first grade and 10 in each of the remaining grades. The children were selected by choosing from the school register those whose ages were nearest to a given age plus six months. Thus, the average age of the first-grade children was 6.8 yr., the second-grade children 7.6 yr., and so forth. The 10 college students were volunteers from classes in elementary psychology. All *Os* were free from visual defects. There were equal numbers of boys and girls in each group.

Procedure. The test-objects were 1-in. wooden dowels, painted flat black (reflect-

---

* Received for publication December 13, 1965. This research was supported by Grant MH-08061 from the Public Health Service. The authors are indebted to the State College School authorities for their coöperation in making subjects available.

[1] H. W. Leibowitz and R. A. Dato, Visual size-constancy as a function of distance for temporarily and permanently monocular observers, this JOURNAL, 79, 1966, 279-284.

[2] E. L. Chalmers, Jr., Monocular and binocular cues in the perception of size and distance, this JOURNAL, 65, 1952, 415-423; A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37.

[3] D. W. Taylor and E. G. Boring, Apparent visual size as a function of distance for monocular observers, this JOURNAL, 55, 1942, 102-105.

[4] H. P. Zeigler and H. W. Leibowitz, Apparent visual size as a function of distance for children and adults, this JOURNAL, 70, 1957, 106-109.

ance 2%) mounted in the center of bases 3½ by 5⅞ by 1⅝-in. The five test-objects were 1.5, 3, 6, 12, and 24-in. high. These were so placed, at distances of 12.5, 25, 50, 100, and 200 ft., that the visual angle subtended by each test-object (height) was constant at 0.57°. *O*'s task was to indicate whether the test-object or a similar comparison-object was "bigger." The series of 43 comparison-objects ranged in height from ½-in. to 48-in., in ½-in. steps to 6-in., and in 1-in. steps thereafter, and were viewed individually 25 ft. and 90° to the right of *O*. Testing took place on the playground of the school at which the children were enrolled.

The children were called for and escorted from their classroom by *E*, who asked each child whether he or she would "like to play a game with us." An effort was made throughout the testing procedure to establish rapport with the child by engaging him in conversation. The task, as explained to *O*, was simply to indicate which of the two sticks was "bigger." No time-limit was placed on *O*. Testing was completed in approximately 20 min.

Half of the *O*s in each class were tested binocularly, the other half monocularly, on a random basis. The *O*s tested monocularly were asked first to look through a tube to determine which eye was dominant. The nondominant eye then was covered during testing.

The testing procedure was first to present a comparison-stimulus which was clearly larger or smaller than the test-stimulus, then another stimulus chosen to elicit the opposite response from the previous one. This procedure was repeated several times until it was clear that the *O* understood the instructions, after which the difference between stimuli was systematically decreased. With this procedure it was possible to determine the size of the comparison-stimulus at which the transition from smaller to larger occurred and from which thresholds were calculated. The response "equal" was allowed when offered.

EXPERIMENT II

Method. The use of test-objects of constant thickness might possibly provide an extraneous cue, since the shape of such a test-object is not independent of its size. Although the instructions stressed that height was the relevant variable, there exists the possibility, especially among young children, that the task might be misinterpreted. To be certain that the effects obtained are not influenced by the nature of the test-object, a control study was carried out with the same procedure, instructions, outdoor environment, and subject-population, but utilizing circular test-objects. These test-objects were so constructed that not only the circular portion of the test-object itself, but also the frame and base on which it was mounted were proportional to viewing distance.

Five disks, 1.5, 3.0, 6.0, 12.0, and 24.0-in. in diameter, were cut from ¼-in. masonite and painted flat white (reflectance, 73.5%). They were mounted on vertical wooden supports (painted flat black, reflectance 4.9%) with their centers 48-in from the ground. The width of the vertical supports was one half the diameter of the disks. The base of the test-object assembly was a rectangular piece of wood 1⅜-in. thick, painted flat black. The longer side of the base, oriented perpendicularly to *O*'s line of vision, was two disk-diameters in length, while the shorter side, parallel to *O*'s line of vision, was equal in length to the diameter of the disk. Thus, all portions of each test-object subtended the same visual angle from *O*'s position.

Observers. The *O*s were 20 kindergarten-children and 10 sixth-grade children, none of whom had served in the previous study. They were tested with half of each group observing monocularly (preferred eye) and half binocularly, on a random basis. The comparison-stimuli consisted of a series of circles cut from white cardboard (reflectance 72.2%) ranging in size from ½ in. to 34 in. in diameter, in ½-in. steps up to 7 in., and 1-in. steps thereafter. These circles were mounted, temporarily, with 'burr' tape on a 4 x 8-ft. background covered with black flock paper (reflectance 0.85%), and viewed at a distance of 25 ft., 90° right of *O*'s line of vision to the test-objects. The centers of the comparison-disks were the same distance from the ground as those of the test-objects (48 in.). The *O*s were asked to point to that one of the two circles—the test-object or the comparison-stimulus—which was 'bigger.' In essence, except for the nature of the test-object and the fact that circular test-objects were mounted above the ground, all conditions and parameters were the same as in Experiment I.

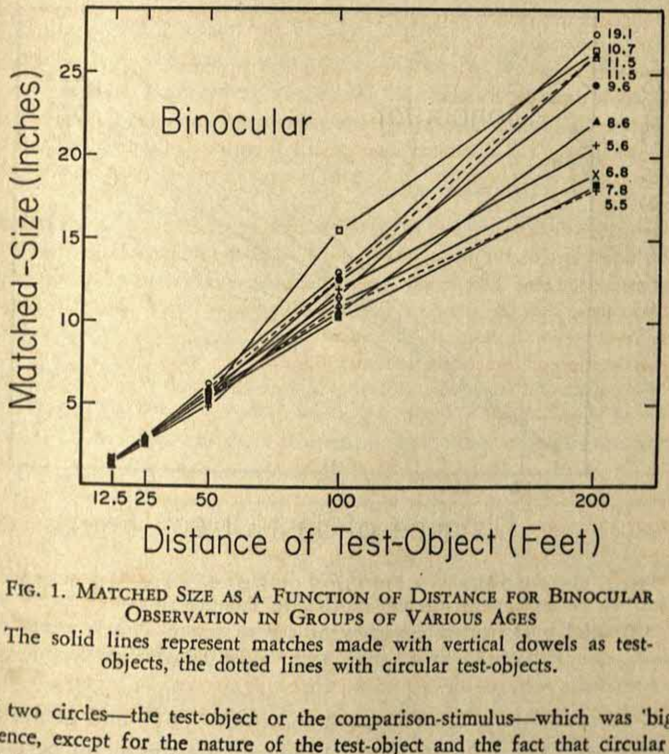


FIG. 1. MATCHED SIZE AS A FUNCTION OF DISTANCE FOR BINOCULAR OBSERVATION IN GROUPS OF VARIOUS AGES

The solid lines represent matches made with vertical dowels as test-objects, the dotted lines with circular test-objects.

RESULTS

Since the matches for the two test-objects were identical for both binocular ($F = 1.18$, $df = 1/29$, $P > 0.25$) and monocular ($F < 1$, $df = 1/29$, $P > 0.25$) observation, the data from both experiments, presented as matched-size in relation to distance with age as parameter, are plotted in Fig. 1 for binocular and in Fig. 2 for monocular observation. The matches for the various age-groups in both graphs are essentially identical for the two closer observation-distances. With increasing distance, however, the matches tend to become systematically lower for the younger *O*s. For the adults, the monocular and binocular observations do not differ ($F < 1$), but for the *O*s of the younger age-groups the monocular matches are generally lower than those for binocular observation. An analysis of variance indicates that the variables of distance ($F =$ 2682.1, $df = 4/336$), age ($F = 12.26$, $df = 7/84$) and monocularity *vs.* binocularity ($F = 14.36$, $df = 1/84$) all are significant beyond the 0.001-level.

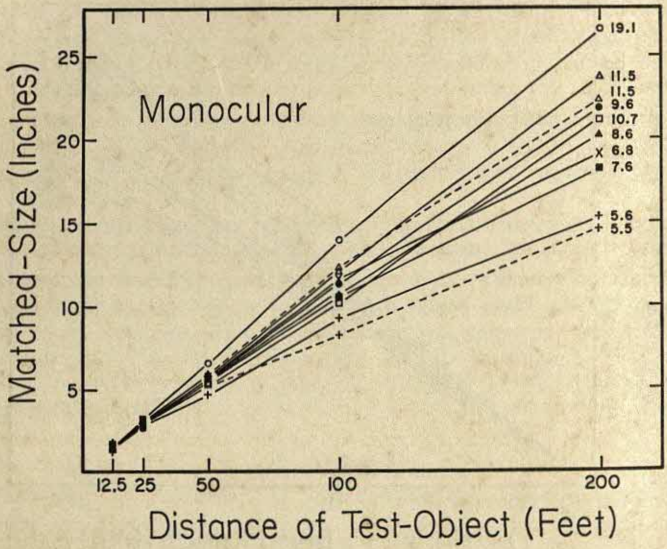


FIG. 2. MATCHED SIZE AS A FUNCTION OF DISTANCE FOR MONOCULAR OBSERVATION IN GROUPS OF VARIOUS AGES

The solid lines represent matches made with vertical dowels as test-objects, the dotted lines with circular test-objects.

Because of the large number of data-points, the theoretical prediction based on the law of size-constancy has not been plotted on the graphs, but the line representing the law of size constancy for the binocular observation would fall approximately on the curve for the 9.6-yr. group, and for monocular observation on the curve for the 11.5-yr. group. Thus, the adults in both situations exhibited a tendency toward 'over-constancy,' as did the 11.5-yr. and 10.7-yr. groups for binocular observation.

Discussion. The results of the present experiment confirm the results obtained in two previous studies which indicated that (1) the tendency toward size-constancy for distant objects decreases with age,[5] and (2) there is no difference between binocular and monocular matches made by adults.[6] It also confirms the hypothesis that the loss of binocular cues is critical only under conditions in which size-constancy is incomplete or marginal. On the basis of these data, it is assumed that the binocular cues are redundant for adult *O*s, for whom size-constancy is essentially completely developed. For the younger *O*s, however, temporary loss of binocular cues results in a reduction of size-constancy.

It should be noted that the matches obtained in the previous study for 8-yr. old children were lower in general than those obtained in the present study for the same age-group with the same procedure and test-objects.[7] This difference may be attributable to the fact that the previous observations were made indoors, while the present experiment was carried out in an outdoor environment. The matches for the adults in the present study are also higher in general than those obtained in several previous studies, all of which were made indoors with the same apparatus and procedure.[8] Apparently, matches tend to be higher in an outdoor as compared to an indoor environment.

The results of the present study emphasize that monocular cues simultaneously present in the visual field are important determiners of size-constancy. This statement does not imply that other mechanisms may not be operative in different situations. Indeed, at distances of a meter or less, cues from the oculomotor adjustment can play a significant role.[9] With familiar test-objects having a known and fixed size, furthermore, cognitive and mnemonic factors undoubtedly are important.[10] Beyond approximately 10 ft. and for test-objects such as those of the present study which can exhibit a wide range of sizes, the available data point, however, to the importance of monocular cues. The critical role played by simultaneous observation of the entire visual field has been well established by a number of studies in which elimination of these cues by a reduction-screen reduces the tendency toward size-constancy.[11] The assumed role of monocular cues also is consistent with the results of studies of the effect of age on one of the monocular cues, *e.g.* perspective, which has been shown to increase systematically as a function of age.[12] The inferior performance of children in size-matching would seem to be a result of their inability to utilize fully those monocular cues which are available.

---

[5] Zeigler and Leibowitz, *op. cit.*, 108.

[6] Leibowitz and Dato, *op. cit.*, 280.

[7] Zeigler and Leibowitz, *op. cit.*, 108.

[8] Leibowitz and Dato, *op. cit.*, 280; Leibowitz, Apparent visual size as a function of distance for mentally deficient subjects, this JOURNAL, 74, 1961, 98-100; H. W. Leibowitz and Vladimir Pishkin, Perceptual size constancy in chronic schizophrenia, *J. consult. Psychol.*, 25, 1961, 196-199.

[9] H. W. Leibowitz and D. Moore, Role of changes in accommodation and convergence in the perception of size, *J. opt. soc. Amer.*, 56, 1966, 1120-1123.

[10] R. C. Bolles and D. E. Bailey, Importance of object recognition in size constancy, *J. exp. Psychol.*, 51, 1956, 222-225; H. W. Leibowitz and L. O. Harvey, Size matching as a function of instructions in a naturalistic environment, *J. exp. Psychol.* (in press).

---

[11] B. E. Holaday, Die Grössenkonstanz der Sehdinge bei Variation der inneren und äussern Wahnehmungsbedingungen, *Arch. ges. Psychol.*, 88, 1933, 419-486; Holway and Boring, *op. cit.*, 29-37.

[12] H. W. Leibowitz and J. M. Judisch, The relation between age and the magnitude of the Ponzo illusion, this JOURNAL, 80, 1967, 105-109; C. Hanley and D. J. Zerbolio, Development changes in five illusions measured by the up-and-down method, *Child Development*, 36, 1965, 437-452; H. W. Leibowitz and M. A. Heisel, L'evolution de l'illusion de Ponzo en fonction de l'âge, *Arch. Psychol.* (Genève), 36, 1958, 328-331.

RESISTANCE TO EXTINCTION IN THE GOLDFISH AS A FUNCTION OF FREQUENCY AND AMOUNT OF REWARD

By R. C. Gonzalez, Nan K. Holmes, and M. E. Bitterman,
Bryn Mawr College

The results of early experiments on partial reinforcement (PRE) in the fish suggested that fish and rat might be differently affected by inconsistency of reward.[1] More recent results suggest that fish and rat perform differently in such experiments because they are differently affected, not by inconsistency of reward, but by frequency and by amount of reward. A difference in the effect of frequency of reward is indicated by the fact that the fish shows the PRE more readily in experiments with equated reinforcements than in experiments with equated trials—when trials are equated, the partial *S*s have fewer reinforcements—while the rat seems to show the PRE as readily in either kind of experiment.[2] Evidence of a difference in the effect of amount of reward comes from experiments on the PRE as a function of amount of reward. While the resistance to extinction of consistently reinforced fish was found in one such experiment to increase with amount of reward,[3] quite the opposite relation has appeared in several experiments with rats.[4] The primary purpose of the present work was to provide some further information on resistance to extinction as a function of amount and frequency of reward in consistently reinforced goldfish, but some supplementary data were collected also on

---

* Received for publication December 17, 1966. This research was supported by Grants MH-02857 and MH-9577 from the United States Public Health Service.

[1] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this Journal, 72, 1959, 184-199; Wodinsky and Bitterman, Resistance to extinction in the fish after extensive training with partial reinforcement, this Journal, 73, 1960, 429-434.

[2] R. C. Gonzalez, R. M. Eskin, and M. E. Bitterman, Extinction in the fish after partial and consistent reinforcement with number of reinforcements equated, *J. comp. physiol. Psychol.*, 55, 1962, 381-386; Further experiments on partial reinforcement in the fish, this Journal, 76, 1963, 366-375.

[3] Gonzalez and Bitterman, The partial reinforcement effect in the goldfish as a function of amount of reward, *J. comp. physiol. Psychol.*, 62, 1967, 342-344.

[4] S. H. Hulse, Amount and percentage of reinforcement and duration of goal confinement in conditioning and extinction, *J. exp. Psychol.*, 56, 1958, 48-57; A. R. Wagner, Effects of amount and percentage of reinforcement and number of acquisition trials on conditioning and extinction, *J. exp. Psychol.*, 62, 1961, 234-242; R. C. Gonzalez, W. A. Roberts, and M. E. Bitterman, Resistance to extinction in normal and extensively decorticated rats as a function of amount and percentage of reinforcement, this Journal, 79, 1966, 542-550.

resistance to extinction as a function of amount of reward in consistently reinforced rats.

METHOD

Subjects. The *S*s were 48 goldfish, 4-5 in. long, and 24 white rats, 90-100 days of age. The animals, all experimentally naïve, were obtained from local dealers. The goldfish were housed individually in 2-gal. tanks and maintained on a 24-hr. feeding schedule of live and dry food. The rats, also housed individually and maintained on a 24-hr. schedule, were reduced to 85% of satiated bodily weight.

Apparatus. The fish were trained in their living tanks, which were carried to a black Plexiglas enclosure diffusely illuminated by white light. The manipulandum was a target of wire mesh, illuminated by red light, at which the *S*s learned to strike. The reward was 1 or 10 *Tubifex* worms discharged from a medicine dropper by a pair of solenoid-operated jaws; as the reward was introduced, the red-target-light was turned off for 1 sec., although there were, of course, other cues associated with the operation of the feeder.

The rats were trained in ventilated picnic chests diffusely illuminated by white light. The manipulandum was a lever, also illuminated by white light. The reward was 1 or 7 Noyes pellets (45-mg.) delivered by a Gerbrands feeder; the lever-light was turned off for 1 sec. whenever reward was introduced, although in this case, too, there were other cues associated with the operation of the feeder.

Both situations were fully automated, and the responses of the animals were registered on cumulative recorders and printing counters.

Procedure. After establishment of the instrumental response by a combination of baiting and shaping methods, there were 15 daily sessions in which the animals were trained on a *CRF* schedule. Four groups of fish ($N = 12$ per group) were given either 10 or 20 reinforcements per session and either 1 or 10 worms per reinforcement in a 2 x 2 factorial design. The total daily ration of each fish was 202 worms and 0.05 gm. of dry food, the portion not earned in the experimental situation being given in the home-tank 1 hr. after each daily session.

Two group of rats ($N = 12$ per group) had 20 reinforcements per session with amount of reward varied. The number of 45-mg. pellets per reinforcement was 1 for one group and 7 for the other. In the sixteenth experimental session, each animal was extinguished for 30 min. On subsequent days, single retraining sessions (in which each group was treated exactly as it had been in the original training sessions) were alternated regularly with 12-min. extinction-sessions. The rats had eight retraining sessions and eight extinction-sessions (in 16 days); the fish had 16 sessions of each kind (in 32 days).

RESULTS

The course of the first extinction in the four groups of goldfish is plotted in Fig. 1 in terms of mean number of responses per block of 5 min. The curves show that the level of response in extinction increases both with frequency and with amount of reward in training. Analysis of variance yields a significant Frequency-effect ($F = 4.07$, $df = 1/44$, $p = 0.05$), a significant Amount-effect ($F = 17.70$, $df = 1/44$, $p <$

0.01), and a significant Blocks-effect ($F = 14.23$, $df = 5/220$, $p < 0.01$), but there are no significant interactions. That the differences in resistance to extinction due to frequency and to amount are highly stable is shown by Fig. 2, in which the data for the subsequent extinctions are plotted. Again for these data, analysis of variance yields a significant Frequency-effect ($F = 4.71$, $df = 1/43$, $p < 0.05$), a significant Amount-effect ($F = 21.68$, $df = 1/43$, $p < 0.01$), as well as a significant Extinction-effect ($F = 8.89$, $df = 15/645$, $p < 0.01$), with no significant interactions. (As the *df*s indicate, one *S* was lost after the initial extinction.)

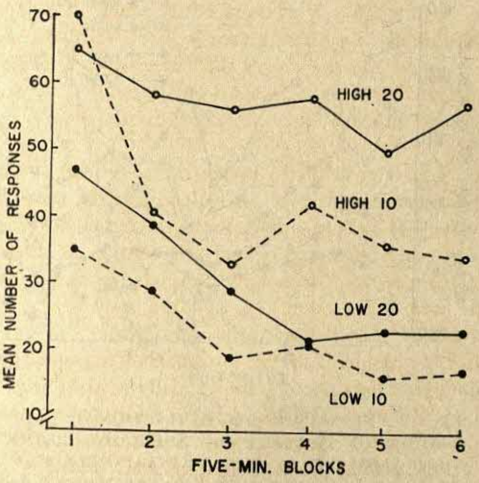


FIG. 1. RESISTANCE TO EXTINCTION IN FOUR GROUPS OF GOLDFISH AFTER 12 SESSIONS OF TRAINING WITH 10 OR 20 CONTINUOUS REINFORCEMENTS PER SESSION AND HIGH OR LOW REWARD

The course of the first extinction in the two groups of rats is plotted in Fig. 3 in terms of mean number of responses per block of 5 min. In the rats as in the fish, the initial level of response is greater for high reward than for low reward, but the two curves soon cross, the level of response in the high-reward group falling below that in the low-reward group. For the rat, then, resistance to extinction is inversely related to amount of reward. Analysis of variance yields no significant Amount-effect ($F < 1$), but a significant Blocks-effect ($F = 68.37$, $df = 5/110$, $p < 0.01$), and a significant interaction ($F = 3.23$, $df = 5/110$, $p < 0.01$). Very much the same pattern appears in the second extinction, after which the difference between the groups disappears as the level of response in both becomes very low.

DISCUSSION

The substantially greater resistance shown in the first extinction by the 20-trial fish as compared with the 10-trial fish was expected from previous experiments on partial reinforcement in the fish. It is easy to understand in the light of these results why, for example, the outcome of an experiment on partial reinforcement in which the partial (50%) group has 20 trials and 10 reinforcements per session should depend on whether the consistent group is trained with equated trials (20 trials and 20 reinforcements per session) or with equated reinforcements (10 trials and 10 reinforcements per session). That number of reinforcements is at least as important with high reward as with low may not seem to fit the fact that when reward is high the fish shows the PRE even with equated trials, but there is no contradiction; high reward increases the resistance of partially reinforced fish much more than it does the resistance of consistently reinforced fish.[5]

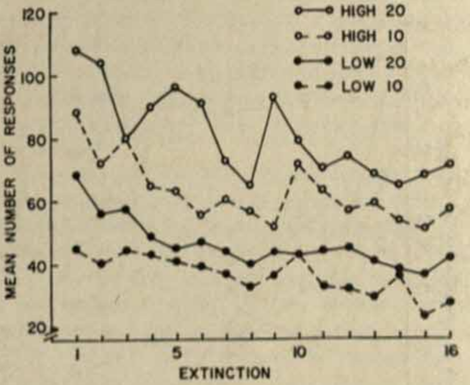


FIG. 2. RESISTANCE TO EXTINCTION IN FOUR GROUPS OF GOLDFISH IN SUCCESSIVE EXTINCTIONS SEPARATED BY SINGLE RETRAINING SESSIONS OF 10 OR 20 CONTINUOUS REINFORCEMENTS AND HIGH OR LOW REWARD

The persistence of the effects of frequency of reinforcement in successive extinctions, and the indication in Fig. 2 of different asymptotic levels for the curves of different frequency, also were to be expected from the results of earlier experiments, in which, after resistance had fallen to asymptote, the amount of interpolated retraining was varied systematically.[6] In those experiments, with each *S* serving as its own control, the fish's resistance to extinction was found to increase with amount of interpolated retraining. The implication of that finding was that independent groups would yield parallel functions of the kind plotted in Fig. 2.

---

[5] Gonzalez and Bitterman, *op. cit.*, 342-344.

The substantially greater resistance shown in the first extinction by high-reward as compared with low-reward groups of fish confirms for another situation (free-operant rather than discrete-trials) and for another measure of resistance (number of responses in a fixed period rather than trials to criterion) the results of an earlier experiment.[7] The present results go beyond the earlier ones also in establishing the persistence of the effect over successive extinctions. They indicate that asymptotic resistance to extinction in the fish is determined by amount as well as by frequency

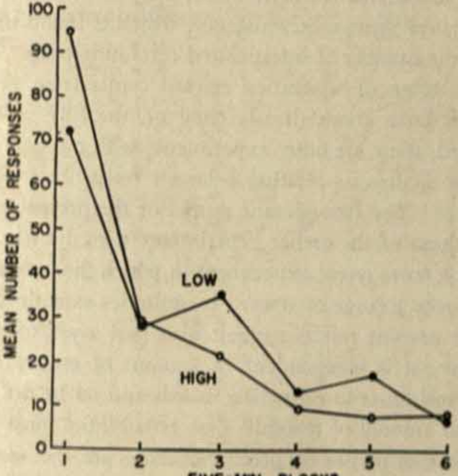


FIG. 3. RESISTANCE TO EXTINCTION IN TWO GROUPS OF RATS AFTER 15 SESSIONS OF TRAINING WITH 20 CONTINUOUS REINFORCEMENTS PER SESSION AND HIGH OR LOW REWARD

of reward. In fact, the present data suggest that amount of reward is the more powerful variable.

---

[6] Gonzalez, Eskin, and Bitterman, Alternating and random partial reinforcement in the fish with some observations on asymptotic resistance to extinction, this JOURNAL, 74, 1961, 561-568; R. C. Gonzalez, N. K. Holmes, and M. E. Bitterman, Asymptotic resistance to extinction in fish and rat as a function of interpolated retraining, *J. comp. physiol. Psychol.*, 63, 1967, 342-344.

[7] Gonzalez and Bitterman, *op. cit.*, 342-344.

The results of this experiment, taken together with those of the previous experiments on partial reinforcement and on asymptotic resistance to extinction, suggest that resistance to extinction in the rat is differently affected than it is in the fish by frequency and amount of reinforcement. Some indication of a difference in the effect of frequency is found in the fact that with low reward the rat seems to show the PRE as readily in equated trials as in equated reinforcements, but the clearest evidence comes from experiments on asymptotic resistance, which is found in the rat to be independent of the number of interpolated retraining trials.[8] Evidence of a difference in the effect of amount of reward comes from a variety of experiments. Apart from those already cited on the PRE as a function of amount of reward, there are other experiments with consistently reinforced rats which show an inverse relation between resistance to extinction and amount of reward.[9] The free-operant results of the present experiment are in accord with those of the earlier experiments with discrete trials, as well as with those of a more recent experiment in which the inverse relation was shown to hold over a range of reward magnitudes extending from 0.045–1.62 gm.[10] The present results suggest also that asymptotic resistance to extinction in the rat is independent of amount of reward.

Why should resistance to extinction in fish and rat be differently related to frequency and amount of reward? Two possibilities must be considered. One is that the mechanism of habit-formation are the same in the two animals, but that the mechanisms of extinction are different. Another is that the mechanisms of habit-formation also are different. It is interesting to note that the results for the fish are precisely what would have been expected on the theory (no longer tenable for the rat) that habit strength is determined by amount and frequency of reinforcement, and that resistance to extinction is determined by habit-strength.

SUMMARY

Goldfish and rats were trained with continuous reinforcement in free-operant situations. There were 15 training sessions followed by a 30-min. extinction-session, after which single retraining sessions were alternated

---

[8] Gonzalez, Eskin, and Bitterman, this JOURNAL, 74, 1961, 561-568; Gonzalez, Holmes, and Bitterman, *op. cit.*, 342-344.

[9] H. L. Armus, Effect of magnitude of reinforcement on acquisition and extinction of a running response, *J. exp. Psychol.*, 58, 1959, 61-63; J. R. Ison and P. E. Cook, Extinction performance as a function of incentive magnitude and number of acquisition trials, *Psychon. Sci.*, 1, 1964, 245-246.

[10] Gonzalez and Bitterman, unpublished data.

with 12-min. extinction-sessions. For four groups of goldfish, number of reinforcements per session (10 or 20) and amount of reinforcement (1 or 10 worms) were varied in factorial design. For two groups of rats given 20 reinforcements per session, amount of reinforcement was varied (1 or 7 pellets). In the fish, resistance to extinction was positively related both to frequency and to amount of reinforcement, not only in the first extinction but throughout the entire series of 16 subsequent extinctions. In the rat, a transient inverse relation between resistance to extinction and amount of reinforcement was found. These and other results suggest that fish and rat are differently affected by the two variables.

MEASUREMENT OF SEQUENTIAL BEHAVIOR BY SUCCESSIVE DIFFERENCES

By ESTELLE SINGER and THORNTON B. ROBY, Tufts University

The optimal strategy in the two-choice probabilistic situation, in terms of objective payoff, is to choose the more likely alternative on every trial. Behavior, however, is generally far from this ideal.[1] It is possible that *Ss* attempt to make their predictions as sensitive as possible to the fluctuations of the series, rather than to optimize over-all payoff in points or money. If so, their predictions would take account of more subtle aspects of the series than over-all frequency; their strategies would be correspondingly more complex.

The present paper is addressed to the problem of how to measure more subtle aspects of binary sequences. Methods that have given meaningful results include counting the number of alternations (or repetitions) and counting the number of runs, which is mathematically equal to the number of alternations plus one.[2] The present study attempted to extend the concepts of repetition and alternation, and to provide a first look at whether the new concepts have empirical relevance.

A repetition occurs in a binary series when two successive choices are the same. To extend this notion to a higher level of abstraction, a higher level repetition will be said to occur when a repeat itself repeats, or when an alternation is followed by an alternation. The number of repetitions of repeating and alternating is a measure at this level analogous to the number of repetitions at the previous, next lower level.

The group of measures being considered thus forms a hierarchy based on the transitions between successive items at each level. An illustrative series of 15 binary responses is presented in Column D_0 of Table I. The measure D_0 is the proportion of 1s in the series and is given on the last line of Column D_0. Column D_1 consists of the absolute value of the difference between adjacent items in Column D_0; Colunm D_2, of the absolute value of the differences between adjacent entries in

---

* Received for publication March 8, 1966. Supported in part by grant G-10947 from the National Science Foundation, and in part by Contract AF 19(628)2450 with the Decision Sciences Laboratory, U.S. Air Force. Computer time was provided by the M.I.T. Computation Center. Michael Feinberg assisted in running the subjects.

[1] G. S. Tune, Response preferences: A review of some relevant literature, *Psychol. Bull.*, 61, 1964, 286-302.

[2] Frank Restle, *Psychology of Judgement and Choice: A Theoretical Essay*, 1961, 107-123.

Column D_1; Column D_3, of the differences between adjacent entries in Column D_2. The measures, D_1, D_2 and D_3, given on the last line, are the proportions of 1s in each Column.

If 1 and 0 are taken to represent responses, then D_0 represents the proportion of times a response of 1 occurred; D_1, the tendency to alternate rather than to repeat responses; D_2, the tendency to change between alternating and repeating, and so on.

TABLE I

ILLUSTRATIVE SERIES

	D_0	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8
	1								
		0							
	1		1						
		1		1					
	0		0		1				
		1		0		0			
	1		0		1		1		
		1		1		1		0	
	0		1		0		1		0
		0		1		0		0	
	0		0		0		1		1
		0		1		1		1	
	0		1		1		0		0
		1		0		1		1	
	1		1		0		1		0
		0		0		0		1	
	1		1		0		0		1
		1		0		0		0	
	0		1		0		0		1
		0		0		0		1	
	0		1		0		1		0
		1		0		1		1	
	1		1		1		0		
		0		1		1			
	1		0		0				
		0		1					
	1		1						
		1							
	0								
Proportions:	.533	.500	.692	.500	.364	.500	.556	.625	.429

If the responses were random and equally probable, the proportions would be 0.50 on the average at all D levels.

The individual responses at a given level are logically determined by the individual responses at lower D levels. The purpose of constructing higher level sequences is to summarize information about the pattern of responses comprising the original series. For example, a D_0 proportion of 0.50 means that 50% of the responses are of each type, but does not indicate the proportion of repetitions (D_1) or the proportion of changes between repetition and alternation (D_2).

The primary purpose of the present study was to determine whether the *S*s are sensitive to information at D levels higher than D_0 (simple response preferences) and D_1 (alternation vs. repetition). It was set up as a two-choice prediction task in which sensitivity at levels up to D_7 was assessed. A secondary purpose was to evaluate the role of immediate memory in determining sensitivity at various D levels.

Method: (1) Subjects. The 112 men and women undergraduate students were paid by the hour, plus a small bonus based on performance. None had previously participated in this type of experiment.

(2) Design. The task for *S* was to predict which figure would be shown next, a star or a square. Each *S* was given four presentation series of 50 trials.[3] Each series had an unusually low proportion (0.32 ± 0.02) at one D level and was neutral (0.50 ± 0.05) at the other levels. Series D_1 had its low proportion at D_1; series D_3 had its low proportion at D_3; series D_5 was low at level D_5; and series D_7 was low at D_7. The *Ss* were expected to alter their response sequences so that the proportions would be low at the appropriate D level.

Four groups of *Ss* were used, each with four series in a different order. The groups were divided in half to assess the role of immediate memory. For half the *Ss* (items available), the correct figures were recorded and left on a blackboard; for the other half (memory), the figures were shown briefly and then erased. The first group therefore did not have to rely on memory for knowledge of prior trials while the second group did. The *Ss* were tested in groups of 7–10. (To provide equal groups in the analysis, data from only 7 randomly selected *Ss* from each session were used.) To prevent confounding of idiosyncracies of administration with test-condition, each set of the eight tests was administered twice. The 16 sessions were given in an approximately counterbalanced order:

(3) Procedure. All *Ss* were treated alike, except as described above. Upon being seated, the *Ss* were read the following instructions.

> You will be asked to predict which figure, a square or a star [*E* writes them on board] will appear next in a long series. Please use the following procedure in recording your predictions. Write "1" and your first prediction on the pad of paper. Then turn the page and wait for the first figure to be written on the board. At the signal, then write "2" and your prediction for the second figure. Continue the same procedure for the rest of the series. Your bonus on this task will be based on the number of correct predictions in excess of chance. We will start with a practice series of eight [eight practice trials]. That was a practice series. We will now begin the main series of 50 items. Please make your first prediction [25 trials]. We will now begin the second half of the series [25 trials more]. That was the end of the first series. We will now begin the second series. Please write "1" on your pad and your prediction.

The same procedure was used until the four series were completed. About 7.0 sec. were allowed for the response and 5.0 secs. for presentation of the correct figure, with one minute between series.

Results. Since the predictions concerned the proportions at the various D levels of the response sequences, these proportions were calculated for each half of the four presentations series. They were then converted by an arc since transformation preparatory to further analysis. A separate analysis of variance was performed at each of four response D levels; D_1,

---

[3] "Presentation series" refers to the sequence of correct answers in a prediction task.

D_3, D_5, and D_7. The mean proportions (untransformed) are shown in Table II.

The main hypothesis predicted lower proportions of response at the D level appropriate to the particular series presented. The decrease could take the form of a lower over-all proportion or of a decrease between the first and second half of the series. Table II shows that the D_1 proportions were lowest for the D_1 series, and that the D_3 proportions were similarly low for the D_3 series. These results proved to be significant at the 0.01 level. Although the over-all proportions for the D_7 series were not low, the proportions for the second half of D_7 were lower than for the first half. This difference was significant at the 0.05 level, but since the initial proportion was greater than 0.50, it is possible that the decrease represented a regression toward the neutral norm of 0.50.

TABLE II

PROPORTION AT EACH D-LEVEL FOR DIFFERENT SERIES
(Each cell represents 25 trials and 112 *S*s.)

Series	Half	Response Level			
		D_1	D_3	D_5	D_7
D_1	1	.451	.485	.487	.520
	2	.440	.475	.480	.500
D_3	1	.486	.470	.483	.480
	2	.470	.462	.504	.517
D_5	1	.486	.496	.486	.498
	2	.458	.506	.493	.506
D_7	1	.459	.490	.487	.517
	2	.450	.487	.498	.505

The type of presentation, included as a variable to assess the influence of immediate memory, was not a factor in the results. Neither the variable itself nor its interactions reached significance.

The residual interaction between series and order was significant for D_3. The result indicated that the *S*s were more sensitive to bias introduced at level D_3 when the biased series occurred at a later time in the experimental session. The results for series D_5 and D_7, while not significant, suggested the same pattern.

The reliability of the individual differences was assessed because of the speculation that individual sensitivity to structural factors in the series in this situation is related to behavior in other situations. The reliabilities were computed for each D level[4] and found to be significant ($p < 0.05$) for the D_1, D_3, and D_7 levels. These results are consistent with the positive results already presented.

Further analyses. In the first additional test of the main hypothesis, that the *S*s could react appropriately to the bias introduced at higher D levels, actual and 'ideal' behavior were compared for the critical D_3 series.[5] Since the proportion at D_3 in the series presented was lower than 0.50, one 'ideal' response would be a proportion of 0.00 at the D_3 level. The response sequence corresponding to such ideal behavior was determined from trial to trial for each *S* and compared with the actual response sequence. The proportion of ideal responses was found to exceed one-half in 69% of the 112 *S*s ($p < 0.01$).

---

[4] E. A. Haggard, *Intraclass Correlation and the Analysis of Variance,* 1958, 79-95.

In a second test, the number of correct predictions was analyzed. Table III shows that the mean proportion correct was greater than 0.50 in each half of each series. Further, the number of *S*s whose proportion correct exceeded 50% was significant ($p < 0.001$) for the second half of each of the four series, and for the first half of series D_1 ($p < 0.001$), D_3 ($p < 0.05$), and D_5 ($p < 0.001$).

It might be suggested that sequences of response were sensitive to attributes of the presentation series because the *S*s responded *merely* by mimicking the outcome of the previous trial. Evidence indicates, however, that minimal mimicry occurred. (1) Mimicry would lead to significant effects of D level for *all* series, (2) The actual proportion of mimicry responses was calculated and found not to exceed chance ($z = 1.38$ by a binomial test). (3) Repeating the outcome of the prior trial could produce more than 50% correct responses only if the series presented had a high D_1 proportion. Yet the *S*s tended to exceed 50% correct in the D_3, D_5, and D_7 series.

TABLE III

PROPORTIONS CORRECT IN VARIOUS CONDITIONS
(Each cell represents 25 trials and 112 *S*s.)

Half	Series			
	D_1	D_3	D_5	D_7
1	.562	.541	.552	.529
2	.615	.557	.570	.538

Discussion. It may be concluded that *S*s learned to respond to a bias in a pseudo-random series at higher levels than D_1. This study therefore supports the view that the *S*s attend to more attributes of the series than the mean relative frequency or the number of repetitions. In addition, since learning was rapid in the sense that the bias was detected and employed within the 50 trials of a series, the *S*s may have been sensitive initially to the characteristics measured by D-level proportions.

The D-level measures may be useful in future research for detecting context effects among adjacent responses and might therefore be explored profitably in situations with less control by the stimulus, such as having predictions made without feedback.

Summary. A group of related measures of binary sequences which represent an extension of the concept of a run was presented. The method employs the notion of successive differences of higher orders. Sensitivity to relevant aspects of sequential patterns was assessed in a standard two-choice prediction-stiuation by introducing a bias in the proportions measured by higher-order differences. It was found that the *S*s did adjust their behavior in a manner appropriate to this bias. Immediate memory appeared to be unrelated to sensitivity to the experimental variable.

---

[5] The D_3 level is critical in that prior research has not directly demonstrated sensitivity at levels higher than D_1.

ESTIMATES OF SIZE AND DISTANCE UNDERWATER

By S. M. LURIA, JO ANN S. KINNEY, and SEYMOUR WEISSMAN, Naval Submarine Medical Center, Groton

Perceived size of an object at different distances usually corresponds to physical size, although the size of the retinal image may vary by a factor of ten or a hundred.[1] According to the size-distance hypothesis, a given visual angle determines a unique ratio of apparent size to apparent distance.[2] If apparent size remains constant while the visual angle decreases, then apparent distance must increase; or, if apparent size increases in proportion to an increase in visual angle, then apparent distance must remain constant.

Underwater viewing provides a unique opportunity to study the size-distance hypothesis because of the change in the size of the retinal image. Light-waves are refracted when passing between media of different density. In going from water to air, they are bent away from the incident-normal, and a virtual image of the submerged object is formed which is displaced toward *O*. Consequently, the retinal image is somewhat magnified, with the degree of magnification increasing as the eye is brought closer to the water.[3]

In the present experiment, estimates of size and distance in air and water were compared. Each *O* compared the size of a distant, submerged target with one presented nearby in air. He also estimated the distance of the submerged target relative to the estimated distance of the near one. The same tasks were performed with all targets in air, with each *O* serving as his own control.

Method. The underwater studies took place in a submerged tower which permitted an underwater view of the targets through a porthole and at the same time provided a place for a comparison-target in air. The facilities also included an underwater track at a depth of 6 ft. on which a small cart could be moved back and forth by a system of pulleys.

---

* Received for publication October 25, 1966. From Bureau of Medicine and Surgery, Navy Department, Research Work Unit MFO11.99-9002. We are indebted to Dr. S. Q. Duntley of the Scripps Institution of Oceanography for the use of the facilities at Diamond Island, N.H., and to Mr. Clarence Pinkham for his assistance during the experiment.

[1] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1955, 480-487; Charles E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 271-285.

[2] See William Epstein, John Park, and Albert Casey, The current status of the size-distance hypotheses, *Psychol. Bull.*, 58, 1961, 491-514.

[3] J. P. C. Southall, *Mirrors, Prisms, and Lenses*, 1923, 95-112.

Judgments of size at distances of 5 and 12 ft. were obtained first. The underwater target, a white 4 x 4-in. square made of metal, was attached to a gray background-plate on the underwater cart. A series of comparison-targets, ranging from 1.5–7.0 in. in ⅛-in. steps, was presented one at a time against a 2-ft. square, dark-gray background on an easel 2 ft. to *O*'s left. Observations were made binocularly with the head in a chin-rest that permitted *O* to swivel his head from side to side. Natural vision, with no attempt to control accommodation, convergence, or pupil-size, was used. The instructions, designed to give 'objective' estimates,[4] were read to him as follows:

> I wish you to estimate the actual size of the square out there. I am going to put a series of squares here (referring to the easel) one at a time and in order, and I wish you to tell me whether these squares are larger or smaller than the one out there. That is, if you actually measured both with a ruler, would this one here be larger or smaller?

The order of presentation was so counterbalanced that half the *O*s made their first estimates in the water while half first observed in air. Moreover, half the *O*s first saw the standard at 5 ft. while half saw it first at 12 ft. Those who first observed the standard at 5 ft. under one condition were shown the standard at 12 ft. under the second condition. The *O*s were randomly assigned to these experimental conditions.

Immediately after giving the size-judgments, the *O*s made distance-estimates. Their task was to estimate the relative distances of a far, comparison target in terms of a near, standard target. The standard was set on the easel 2 ft. from *O*'s chin-rest; this distance constituted a 'standard unit' although *O* was never told that it was actually 2 ft. The comparison-target was positioned randomly at 1-ft. intervals from 4–15 ft. from him. Both standard- and comparison- targets were 4-in. squares. The instructions were as follows:

> With your head in this chin-rest, the distance from you to this card is one standard unit. We want to know how many standard units it is from you to that card out there. We are going to move that card back and forth, and each time you tell us how many standard units it is from you. In making your estimates, you may look back and forth between the two cards. Try not to take too long. It may be difficult at times but just do the best you can. Any questions? Before we begin, make an estimate of this distance.

After *O* had made an estimate, *E* said, "You mean that this distance [pointing to the standard] will go into that distance [pointing to the comparison] that many times?"

In the control study, both the targets and the standard were presented in air; it differed only in that the *O*s were seated out of doors, and the distant target was fixed to a black vertical steel rod and viewed against a large black curtain at a distance of 25 ft. The *O*s were 20 men and women selected at random from those working and vacationing at Lake Winnipesaukee, N.H. They ranged in age from 16-60 yr.

Results. Table I gives the average estimates of the size of the 4-in. target and the *SD*s for the 20 *O*s. At both viewing distances, the average estimates in air are somewhat too small. Nevertheless, the estimates are very close to the physical size and represent considerable size-constancy—estimates in accord with retinal size would be 1.6 in. at 5 ft. and 0.67 in. at 12 ft. A comparison of the results in water with those in air shows an increase in perceived size at both distances. The differences are significant at the 1% level for both viewing distances. The increase in apparent area is about 25% for both distances, while the physical magnification occurring as a result of refraction from water to air is approximately 30% for these conditions.

---

[4] A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192.

The median estimates of distance by the 20 *O*s are shown in Fig. 1 in 'standard units.' Medians rather than means are used, because some *O*s tended to make extremely high estimates of the farther distances which would have disproportionately affected the means. The median estimates are considerably higher for the underwater targets than for those in air except at the nearest distances. This result clearly appears in the data for 14 of the 20 *O*s. Five *O*s showed little difference between the two conditions and one showed a reverse effect. It is of interest, however, that of the 6 *O*s who did not show a tendency to overestimate distances in water, four (including the reversal) were in the group which first observed in air. This fact suggests that the second set of estimates, made in water, was strongly influenced by the first set of estimates made in air.

TABLE I

MEANS AND *SD*s OF ESTIMATES OF THE SIZE (IN IN.) OF THE 4-IN. STANDARD BY 20 *O*s

Medium		Distance from *O* (ft.)	
		5	12
Air	Mean	3.76	3.81
	SD	0.25	0.39
Water	Mean	4.19	4.28
	SD	0.44	0.76
Diff. in area (in.^{2})		3.42	3.60
Percent increase		24.2*	24.8*

* Significant at 1% level.

Discussion. It seems quite clear that the size of unknown objects underwater is overestimated when compared with estimates in air. This result conforms with theoretical expectations regarding the enlargement of the retinal image resulting from refraction of the light-rays when they pass from water to air. Furthermore, the amount of increase, by area, calculated from the refractive indices of air and water is comparable to the increase in perceived size actually observed.

The results of the distance-estimation experiment are, however, surprising. The size-distance hypothesis would lead us to expect, on the basis of the increased size of the retinal image of the underwater target, that the visual perception would be either of a larger target at an accurate distance or a target of correct size at a closer distance (or some intermediate combination). This expectation was not realized. There were instead quite accurate estimates of size (when the increase in the retinal image is taken into account) but overestimates of distance. These results, then, do not fit the size-distance hypothesis. We agree, therefore, with Epstein, Park, and Casey that it cannot be considered a sufficient explanation of space-perception.[5] It is evident, as Gruber noted some time ago,[6]

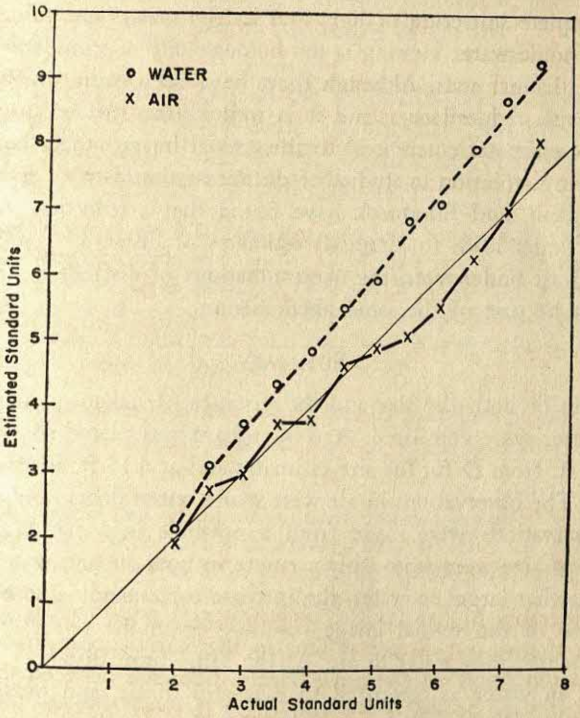


FIG. 1. MEDIAN ESTIMATES OF DISTANCE BY 20 *O*S IN AIR AND WATER
The thin solid line represents the function for correct estimates.

---

[5] Epstein, Park, and Casey, *op. cit.*, 511.

[6] H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-426.

that *O*s make estimates of size which are not related to their estimates of distance. Epstein has suggested that judgments of size may be determined by the absolute size of the retinal image, but that estimates of distance may depend on the relative size of the retinal images of successively presented stimuli.[7]

What is the reason for the overestimations of distance? Due to the almost complete scattering of light-rays by the water,[8] the salient characteristics of underwater viewing is the homogeneity of stimulation and the absence of distinct cues. Although there have been many studies of size-estimates with reduced cues, and it is well known that under such conditions size-estimates correspond to the retinal image,[9] there has been no corresponding attention to studies of distance-estimates with reduced cues. Fry, Bridgman, and Ellerbrock have found that a reduction in the contrast of a target leads to increased estimates of distance.[10] Given the reduced contrast underwater, the overestimations of distance in the present study must be part of the same phenomenon.

SUMMARY

Estimates of both the size and the distance of unknown objects in air and in water were compared. A 4-in. square was placed at distances of 5 and 12 ft. from *O* for the size-estimates and at 4-15 ft. for the distance-estimates. The observations in air were made out-of-doors, and the underwater observations were made from a porthole in a submerged tower. Estimates of size were reasonably accurate in both air and water, but they were somewhat larger in water, the increase corresponding to the increase in the size of the retinal image resulting from the refraction of light waves passing from water to air. At the same time, distances were overestimated in water as compared to estimates in air, and overestimations increased with distance. It is concluded that these results are not explained by size-distance invariance, and that in the absence of distinct cues, distance will be overestimated.

---

[7] William Epstein, The known-size-apparent-distance hypothesis, this JOURNAL, 74, 1961, 333-346.

[8] S. Q. Duntley, Light in the sea, *J. opt. Soc. Amer.*, 53, 1963, 214-233.

[9] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance invariant, this JOURNAL, 54, 1941, 21-37.

[10] G. A. Fry, C. S. Bridgman, and V. J. Ellerbrock, The effects of atmospheric scattering on binocular depth perception, *Amer. J. Optom.*, 26, 1949, 9-15.

A DYNAMIC CUE OF THE MAGNITUDE OF TACTILE SENSATION

By QUANA R. JONES, JR., Auburn University, and
W. W. DAWSON, University of Florida

At least two components exist in the application of a mechanical stimulus to the skin. These are the static (displacement and skin-area) and the dynamic components (velocity and acceleration). Early theories of the tactile sense emphasized the static component.[1]

Publications by Nafe, and Nafe and Kenshalo have proposed the movement of cutaneous nerve-endings as the stimulus for all of the cutaneous qualities.[2] Support for this view is found in the search by Nafe and Wagoner who were unable to relate any static stimulus component to verbal reports to stimulus failure (adaptation) and in Nafe and Kenshalo who described afferent neuron excitation by constant velocity contact stimulation of tissues of rat and frog.[3]

This report demonstrates the perceptual importance of the dynamic component of tactile stimuli of equal and of differing contact area and applied force.

Method: Apparatus. Two sets of six cylinders provided stimuli of systematically varying contact area. Set A exerted a constant force of 5.2 x 10^4 dynes. Set B exerted forces from 2.4 x 10^4 to 6.3 x 10^4 dynes. All cylinders were solid and were machined from brass stock. The base of any cylinder in a set was identical in diameter to the base of the correspondingly numbered cylinder in the other set. In each set the cylinders varied in base diameter by 2.54-mm. increments from 20.32 to 33.02 mm. and were numbered 1 through 6, respectively. The 'standard' stimulus (base 3.63 cm.^{2}) in each set was cylinder No. 1. The remaining 5 cylinders were the 'comparison' stimuli.

Procedure. As diagrammed in Fig. 1 the experimental design was the 2 x 2 factorially randomized block-type. Within the design, the *O*s were assigned in 5 blocks of 4 *O*s each. Within each block, 2 treatments (each of which had 2 levels)

---

* Received for publication September 2, 1966.

[1] E. H. Weber, Der Tastsinn und das Gemeingefuhl, in Rudolph *Wagner's Handworterbuch der Physiologie,* 3, 1846, 481; G. Meissner, Untersuchengen uber den tastsinn, *Z. rat. Med.,* 7, 1859, 92-188; Max von Frey and Friederich Kiesow, Ueber die function der Tastkorperchen, *Z. Psychol.,* 20, 1899, 126-163.

[2] J. P. Nafe, The pressure, pain, and temperature senses, in C. A. Murchison (ed). *A Handbook of General Experimental Psychology,* 1934, 1037-1087; J. P. Nafe and D. R. Kenshalo, A quantitative theory of feeling, *Ann. rev. Psychol.,* 69, 1960, 15-32.

[3] J. P. Nafe and K. S. Wagoner, The nature of pressure adaptation. *J. gen Psychol.,* 25, 1941, 295-321; Nafe and Kenshalo, Stimulation and neural response, this JOURNAL, 71, 1958, 199-208.

were assigned at random with one *O* for each level of treatment, *i.e.* a total of four *O*s per block. Blocks were numbered one through five and corresponded to base diameters of the comparison stimuli from smallest to largest. One treatment consisted of stimulus comparisons in which force was constant, and in the other force was varied systematically. Levels of each treatment were established by order of presentation of the 'standard' stimulus and 'comparison' stimulus. The *O*s, aged 15-18 yr., were 12 men and 8 women highschool students. For each *O* the skin-area above the left knee was shaved 24-hr. prior to his participation in the experiment.

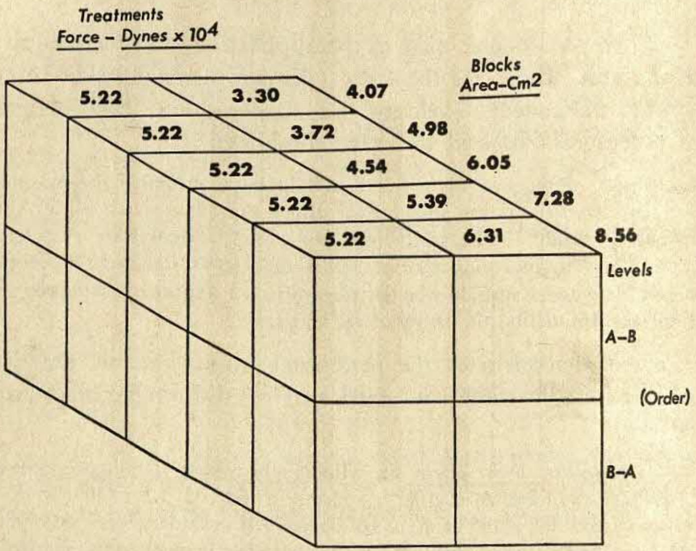


FIG. 1. EXPERIMENTAL DESIGN SHOWING PARAMETERS OF THE COMPARISON-STIMULI

Four *O*s were assigned to each cell. Standard stimuli, not shown, delivered 2.4 x 10^4 dynes for the force variable treatment and 5.22 x 10^4 dynes for the force constant treatment. Both standard base areas were 3.63 cm^2

The stimulus-cylinders were brought to body temperature by having an assistant clasp them tightly in her hands for 15 min. prior to each experimental session.

Before entering the test-room, each *O* was blindfolded. Then he received 10 practice trials followed by 8 trials in which his judgments were recorded. In each trial 2 stimuli ('standard' and 'comparison') were applied consecutively on the skin-surface. After 4 sec. the first stimulus was removed and the other stimulus (with which it was compared) was placed. After 4 sec., the second stimulus was removed and the trial ended. After a 30-sec. interval the next trial began. At the end of each trial (comparison), *O* was required to state whether there was a difference in 'weight' between the presentations of the standard and comparison. Comparison stimulus-positive judgments were recorded and subjected to analysis of variance.

Results. The percentage of positive judgments elicited by constant and variable force comparisons over the several contact-areas is presented by the Fig. 2. Analysis of variance of the raw data (Table I) resulted in identification of the force variable as the only significant factor in the judgment. Bartlett's test for homogeneity applied to these data yielded a small and insignificant X^2 value.

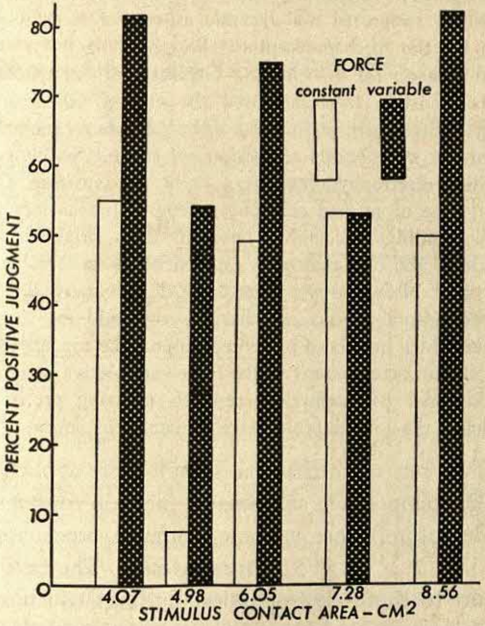


FIG. 2. DISTRIBUTION OF THE *Os*' JUDGMENTS

TABLE I

ANALYSIS OF VARIANCE SUMMARY

Source	Sum of squares	*d.f.*	Mean square	*F*
force	22.05	1	22.05	6.43*
order	0.05	1	0.05	
interaction	2.45	1	2.45	
area	23.20	4	5.80	
error	41.20	12	3.43	
Total	88.95	19		

* $p<0.05$.

A perfect result would have produced 100% positive judgments for the force-variable and chance or below percentages for the constant series. The standard stimuli in each series presented the least area for constant force and also least area for the variable force. No trends in discriminability were apparent for increments in area, force constant, or increments in area and force.

Discussion. Electrophysiological research by Nafe and Kenshalo and by Werner and Mountcastle has supported the dynamic aspects of mechanical stimuli as the adequate stimuli for the mechanoreceptors.[4] Elegant delivery systems used by these authors provided data on the velocity and displacement components of stimulation. Hubbard and Loewenstein have measured the energy delivered to the Pacinian corpuscle during excitation of its afferent fiber.[5] Although inquiries into the relevance of the dynamic components of mechanical stimuli for human cutaneous perception may utilize electrophysiology as a point of departure, the morphological contrast of the tongue of rat and skin of the frog to human hirsute skin limits the generality of the findings of animal research. Data originating in the isolated Pacinian corpuscle of the cat's mesentary is particularly susceptable to this criticism.

Iggo, Tapper, and Nafe and Wagoner have demonstrated the sensitivity of cutaneous mechanoreceptors in cats and humans to rapid movements less than 50 microns.[6] The data which have been presented support the conclusion that the dynamic result (velocity and, or, acceleration) of the force-variable is a major one in tactile discrimination. Additional psychophysical research utilizing precise stimulus-control is required to define the importance of each dynamic component of the stimulus.

Summary. Two sets of mechanical stimuli were applied to the skin of human *O*s. With temperature and area of the skin constant, stimulus-area, force, and order of reference variable, stimulus presentation were varied systematically in a 2 × 2 × 5 factorial design. The results lend psychophysical support to those electrophysical findings which suggest that the dynamic (velocity or acceleration) results of stimulus-force are a major cue for tactile discriminations.

---

[4] Nafe and Kenshalo, *op. cit.*, 1958, 204; G. Werner and V. B. Mountcastle, Neural activity in mechanoreceptive cutaneous afferents: stimulus response relations, Weber functions and information transmission, *J. Neurophysiol.*, 28, 1965, 281.

[5] S. J. Hubbard, A study of rapid mechanical events in a mechano-receptor. *J. Physiol.*, 141, 1958, 198-218; W. R. Loewenstein, Biological transducers, *Sci. Amer.*, 1960, 98-106.

[6] A. Iggo, An electrophysiological analysis of afferent fibers in primate skin. *Acta Neuroveg.*, 24, 1963, 225-240; D. N. Tapper, Input-out relationships of a tactile sensory unit of the cat. *Trans. N.Y. Acad. Sci.*, 26, 1964, 697-701; Nafe and Wagoner, *op. cit.*, 302 f.

RELATIONSHIPS AMONG A NUMBER OF MEASURES OF MEANINGFULNESS

By JOHN F. HALL, The Pennsylvania State University

The search for the determinants of the meaningfulness of discrete verbal units has been extensive with measures of (1) association, single and multiple, (2) pronunciability, (3) subjective frequency or familiarity, and (4) objective frequency, all being examined. What, however, is the relationship among these measures? As Goss and Nodine's review indicates, relatively few investigators have been interested in this question.[1] In the present experiment, the interrelationships among the four listed measures was examined.

METHOD

Materials. The 200 trigrams used were selected from four sources to provide a wide range of experimental material. The four sources, with the number of trigrams selected from each, were as follows: (1) Underwood-Schulz frequency-count, 70; (2) Thorndike-Lorge word-count, 58; (3) Noble's *a′*-values, 48, (4) Witmer's list of consonant-syllables, 24.[2] Trigrams were chosen which sampled the entire range of values associated with each source.

Three 9-point graphic rating-scales were constructed to measure: (1) pronunciability, (2) subjective frequency or familiarity, and (3) number of associations. The pronunciability-scale was similar to that employed by Underwood and Schulz although somewhat different categorical descriptions were used; *i.e.* IMPOSSIBLE, VERY HARD, HARD, BELOW AVERAGE, AVERAGE, ABOVE AVERAGE, EASY, VERY EASY, AND EXTREMELY EASY.[3] The subjective-frequency scale bore some likeness to Noble's familiarity-scale and the Underwood-Schulz judged-frequency scale.[4] With this scale, the *Ss* were instructed to rate verbal items as to the number of times they had seen, heard, or used them in speaking, writing, or reading. Instructions also indicated that some of the trigrams might be found as parts of words. For example, ZLE was used as an example of a trigram to be found in the word PUZZLE. The categories employed were: NEVER, VERY RARELY,

---

[*] Received for publication October 10, 1966.

[1] A. E. Goss and C. F. Nodine, *Paired-Associates Learning*, 1965, 1-358.

[2] B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Learning*, 1960, 1-430; E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944, 1-274. C. E. Noble, Measurements of association value (*a*), rated associations (*a′*), and scaled meaningfulness (*m′*) for the 2100 CVC combinations of the English alphabet, *Psychol. Repts.*, 8, 1961, 487-521; L. R. Witmer, The association value of three-place consonant syllables, *J. genet. Psychol.*, 47, 1935, 337-360.

[3] Underwood and Schulz, *op. cit.*, 19-20.

[4] Noble, The meaning-familiarity relationship, *Psychol. Rev.*, 60, 1953, 89-98; Underwood and Schulz, *op. cit.*, 184-185.

RARELY, BELOW AVERAGE, AVERAGE, ABOVE AVERAGE, OFTEN, VERY OFTEN, and ALMOST ALWAYS. The associative scale, modeled after that used by Noble, Stockwell, and Pryer, required *Ss* to judge the number of associations they had to the trigrams which were presented.[5] The categories employed were: NONE, VERY FEW, FEW, BELOW AVERAGE, AVERAGE, ABOVE AVERAGE, MANY, VERY MANY, and A GREAT MANY. It was possible also to approximate Glaze's measure by noting the relative frequencies of response in categories beyond NONE.[6]

An objective frequency-count for the 200 trigrams was obtained from the Underwood-Schulz (Appendix D) count, with their total count being used.[7] If a given trigram was not found in this list, a value of zero was assigned. Because of the skewness of this distribution, a value of 1 was added to all counts, with these counts then being transformed into logarithms.

Procedure. The *Ss*, 100 students in Introductory Psychology, were given the three rating-scales and asked to provide ratings for each of the trigrams. For each scale, 13 trigrams were printed on a page, and the pages were assembled in random order. The order of rating-scale presentation was counterbalanced, providing six different orders.

TABLE I

CORRELATIONS AMONG MEASURES OF MEANINGFULNESS

	Rated pronunciability			Rated frequency			Rated association			Glaze association		
	Total	Words	Non-words	Total	Words	Non-words	Total	Words	Non-words	Total	Words	Non-words
Rated frequency	.75	.36	.71									
Rated association	.88	.66	.81	.88	.71	.90						
Glaze association	.90	.64	.87	.84	.42	.86	.87	.67	.90			
U-S frequency	.37	.09	.51	.78	.76	.88	.50	.37	.71	.54	.19	.72

Results. Of 60,000 possible responses, 202 were lost because certain *Ss* neglected to provide a rating for particular trigrams. The mean rated value for each trigram for each of the three scales was obtained from the remaining responses; the mean of each trigram was based on a minimum of 97 individual ratings. The reliability of each of these scales was obtained by taking the mean values for the 200 trigrams secured from a randomly selected half ($N = 50$) of the experimental sample and then correlating these values with the corresponding values for the other half. The reliability-coefficients were as follows: Pronunciability = 0.99; Association = 0.97; Subjective Frequency = 0.99.

The correlations obtained among the varying measures of meaningfulness using all of the 200 items can be found in Table I under "Totals." If we examine the trigrams as to whether they were words or non-words, we find 99 of them were words (as found in Webster's Collegiate Dictionary) while 101 were not. The correlations among the varying measures for words and non-words also can be found in Table I.

---

[5] C. E. Noble, F. E. Stockwell, and M. W. Pryer, Meaningfulness (m') and association value (a) in paired-associate syllable learning, *Psychol. Repts.*, 3, 1957, 441-452.

[6] J. A. Glaze, The association value of non-sense syllables, *J. genet. Psychol.*, 35, 1928, 255-269.

[7] Underwood and Schulz, *op. cit.*, 332-369.

The interrelationships among the varying measures, with the exception of the *U-S* frequency-count are quite high. Additional support for the validity of these values is found in the correlations obtained between pronunciability and subjective frequency ($r = 0.73, 0.79$), pronunciability and the *U-S* frequency count ($r = 0.38$), and the *U-S* frequency-count and subjective frequency ($r = 0.46, 0.51$) as reported by Underwood and Schulz.[8]

When we compare the correlation for words and non words, we find lower correlations among the measures for words. An examination of the distributions reveals that restriction of range—especially in the case of pronunciability-ratings—may account for the lowered correlations.

It must be noted that the correlations reported here are undoubtedly spuriously high, since there is some evidence to support the position that when *S* is requested to rate a verbal item on a given attribute, he may use other characteristics of the material upon which to base his rating. For example, on a *priori* grounds, almost all consonant-syllables would appear to be equally difficult to pronounce, yet there is a high correlation between the rated pronunciability of these units and their associative value. Nonetheless, it seems obvious from the magnitude of the correlations obtained, that the measures of pronunciability, subjective frequency, and association have much in common, and that a single measure would satisfy the experimental requirements of many investigators.

---

[8] Underwood and Schulz, *op. cit.,* 161, 186.

SIZE-CONSTANCY IN OLDER PERSONS: A FUNCTION OF DISTANCE

By H. W. LEIBOWITZ and J. M. JUDISCH, Pennsylvania State University

It has been suggested that the Ponzo illusion of perspective is caused by the 'misapplication' of a mechanism which ordinarily subserves the phenomenon of size-constancy.[1] Specifically, it is assumed that *O* learns to associate converging lines with distance in such a way that the nearer an object to the point of convergence, the greater the correction, and the larger it will seem. Support for this hypothesis is provided by the fact that both the magnitude of the Ponzo illusion and the degree of size-constancy for distant objects increase from childhood through adulthood.[2] Recently, however, it has been demonstrated that the Ponzo illusion decreases markedly among the aged.[3] If the mechanisms subserving size-constancy and the Ponzo illusion are indeed identical, size-constancy should also show a decrease as a function of age.

Method. The 19 *O*s ranged in age from 51–88 yr. (mean age, 73.6 yr.). All had served in the previous study of the Ponzo illusion. The apparatus and procedure have been described previously.[4] *O* observed one of a series of white disks at distances of 10, 20, 40, 80, or 120 ft. The comparison-stimulus was provided by a projector which produced a variable-size circular image (luminance, 0.63 ft.-L.) on a screen 20 ft. from and 90° to the right of *O*. *E* adjusted the size of the comparison-stimulus until *O* indicated that it was 'the same size' as the test-object. The luminance of the stimuli was approximately 8 ft.-L. with a variation no greater than ± 4 ft.-L.

Results. Mean matched size as a function of distance is plotted in Fig. 1. For comparison, the matches made by a group of *O*s of college-age

---

* Received for publication October 28, 1966. This research was supported by Grant MH-08061 from the National Institute of Mental Health. The coöperation of the authorities of the Centrecrest Home, Bellefonte, Pa., is gratefully acknowledged.

[1] R. Tausch, Optische Täuschungen als artifizielle Effekte der Gestaltungprozesse von Grössen-und Formenkonstanz in der naturlichen Raumwahrnehmung, *Psych. Forsch,* 24, 1954, 299-348; E. von Holst, Aktive Leistungen der menschlichen Gesichtswahrnehmung, *Studium Generale,* 10, 1957, 231-243; R. L. Gregory, *Eye and Brain,* 1966, 145-156.

[2] H. P. Zeigler and H. W. Leibowitz, Apparent visual size as a function of distance for children and adults, this JOURNAL, 70, 1957, 106-109; H. Leibowitz et M. Heisel, L'evolution de l'illusion de Ponzo en fonction de l'âge, *Arch. Psychol. (Geneva),* 36, 1958, 328-331.

[3] H. W. Leibowitz and J. M. Judisch, The relation between age and the magnitude of the Ponzo illusion, this JOURNAL, 80, 1967, 105-109.

[4] H. W. Leibowitz and R. A. Dato, Visual size-constancy as a function of distance for temporarily and permanently monocular observers, this JOURNAL, 79, 1966, 279-284.

are plotted on the same graph.[5] It will be noted that the points for the shorter distances agree with the prediction based on the 'law' of size-constancy. With increasing distance, the data tend to fall below the theoretical prediction for both groups. The difference between the two groups is not significant ($F = 4.09$, $df = 1/37$, $p > 0.05$). Thus, although the magnitude of the Ponzo illusion decreases in old age, the tendency toward size-constancy remains the same.

Discussion. The question may be raised as to whether the decrease in the Ponzo illusion with age might be attributable to some general factors associated with aging. Relevant evidence is available from a study of the Poggendorff illusion, which exhibits a different functional relation over the same age-range and, in particular, remains at the same level from the age of 10 yr. through that of the oldest *O*s tested.[6] If dioptric or physiological factors were responsible for the decrease in the Ponzo illusion, it would be expected that they would also come into play when viewing the Poggendorff figure. Secondly, the visual acuity of the *O*s in the present study was tested at both 10 and 120 ft. by means of a Landolt 'C' test-object.[7] The Pearson r between age and acuity is -0.38 at 10 ft. ($df =$

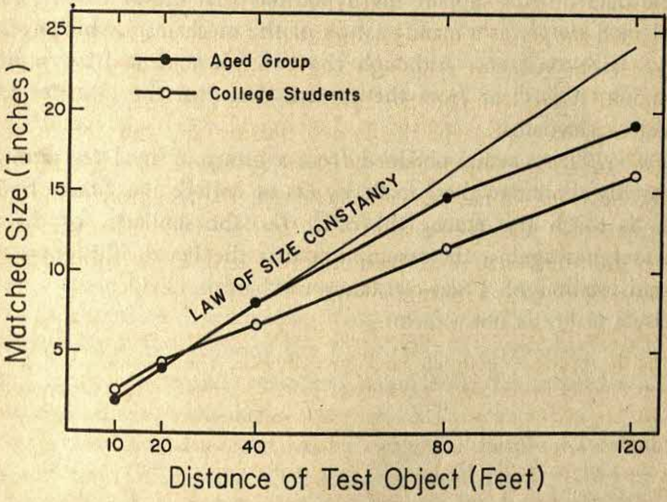


FIG. 1. MATCHED-SIZE AS A FUNCTION OF DISTANCE FOR ELDERLY *O*S AND FOR *O*S OF COLLEGE AGE

---

[5] Leibowitz and Dato, *op. cit.*, 279-284.

[6] H. W. Leibowitz and J. Gwozdecki, The magnitude of the Poggendorff illusion as a function of age, *Child Devel.* (in press).

[7] Because of the marked variability in acuity among the aged (J. E. Birren, *Handbook of Aging and the Individual*, 1959, 511) acuity-data were obtained for the same *O*s who had previously served in the Ponzo and size-matching studies.

17, $p < 0.05$) and —0.56 at 120 ft. ($df = 17$, $p < 0.05$) indicating a tendency for acuity to decrease slightly with age, particularly at the longer distance. This decrease in acuity does not, however, correlate significantly either with size-matching or with the magnitude of the Ponzo illusion. The correlation between age and the size-matches at 120 ft. is —0.24, and between acuity at 10 ft. and the magnitude of the Ponzo illusion is —0.16. Thus, the slightly lower acuity of the older *O*s is not correlated either with their size-matches or with the magnitude of the Ponzo illusion.

These data do not support the hypothesis that the Ponzo illusion can be explained *simply* as a manifestation of the mechanism which ordinarily subserves size-constancy. Although the two phenomena have something in common, it is clear from the present data that they do not rest on identical mechanisms.

Summary. Size-matches obtained from a group of aged *O*s were found to be highly similar to those made by *O*s of college age. Since the Ponzo illusion is much less among the older *O*s, the similarity of their size-matches argues against the assumption that the Ponzo illusion is merely the manifestation of a size-constancy mechanism. Evidence is presented that visual acuity is not a factor.

APPARATUS

EQUIPMENT FOR VARYING THE INTENSITY OF LIGHT

By JAMES P. THOMAS, University of California, Los Angeles

The optical equipment used in most studies of visual perception produces sharply focused images of uniform intensity. Some perceptual phenomena can be observed, however, only with other types of images. For example, Mach bands, which have been the subject of much theorizing, are observed only when contours are blurred or when the intensive change from one area to another is gradual, rather than step-like.[1] As another example, Cornsweet has devised an illusion in which two parts of a field are made to appear different in brightness.[2] The illusion is produced by bisecting the field with a perturbation which consists of both continuous and discontinuous changes in intensity as a function of distance. To study these, and other, phenomena, an apparatus is required which produces a variety of intensive distributions. This article describes principles, and two instruments which embody the principles, for producing continuous and discontinuous intensive variations along both the vertical and horizontal axes of a visual display.

Several techniques have been used to produce continuous and discontinuous variations of intensity. One technique uses rotating disks or cylinders.[3] A second method utilizes diffraction.[4] A third is based upon the fact that varying the area of a source of light, thereby varying the total flux emitted, varies the intensity with which the source illuminates a screen. Shadow casters,[5] slits,[6] and cylindrical lenses[7] can be used to control the area of the source which illuminates different points on the screen

---

[*] This equipment was developed during research supported in part by USPHS Grant NB05185 from the National Institute of Neurological Diseases and Blindness.

[1] C. H. Graham, Visual form perception, in C. H. Graham, (ed.), *Vision and Visual Perception*, 1965, 549-552.

[2] Floyd Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina*, 1965, 74-76.

[3] Graham, *op. cit.*, 549-550.

[4] Gerald Westheimer, Modulation thresholds for sinusoidal light distributions on the retina, *J. Physiol.*, 152, 1960, 67-74.

[5] G. A. Fry, Mechanisms subserving simultaneous brightness contrast, *Am. J. Optom. and Arch. Am. Acad. Optom.*, 25, 1948, 162-178.

[6] Ernst Mach, On the influence of spatially and temporally varying light stimuli on visual perception, in Ratliff, *op. cit.*, 323-324.

[7] Mach, *op. cit.*, 324.

and, thus, the intensity at every point. The apparatus described in this article falls in the third category and uses cylindrical lenses to control the intensive distribution along both axes of the screen.

The principles to be described are illustrated by the simplified optical system shown at the top of Figure 1. The system consists of a small source of light, centered on the vertical *i*, at Position *A;* a small aperture, centered on the horizontal *j*, at Positions *B;* and two cylindrical lenses, L_a and L_b, oriented at 90° with respect to each other. Lens L_a forms images on the horizontal axis and is so placed that, on the horizontal axis, an image of the source is formed at *A′B′*, *i.e.* a vertical strip of light centered on *i′*. Lens L_b forms images on the vertical axis and is so placed that, on the vertical axis, an image of the aperture is formed at *A′B′*, *i.e.* a horizontal strip of light centered on *j′*. Acting together, the two lenses form a complete image, *i.e.* a rectangle of light centered on *i′j′*. The width of this image is determined by the width of the source at *A*, and the height of the image is determined by the height of the aperture at *B*. If the height, *x*, of the source is altered, the size of the image is unaffected, but the intensity of the image is changed. For example, if the height is increased, there are increases in the total amount of light given off by the source, the amount passing through the aperture, and the amount falling in the image. Thus, the intensity of the image is proportional to the height, *x*, of the source. Similarly, changing the width, *y*, of the aperture does not affect the size of the image, but does alter its intensity. As the width is increased, more light from the source is allowed to pass and more light falls in the image. Thus, the intensity of the

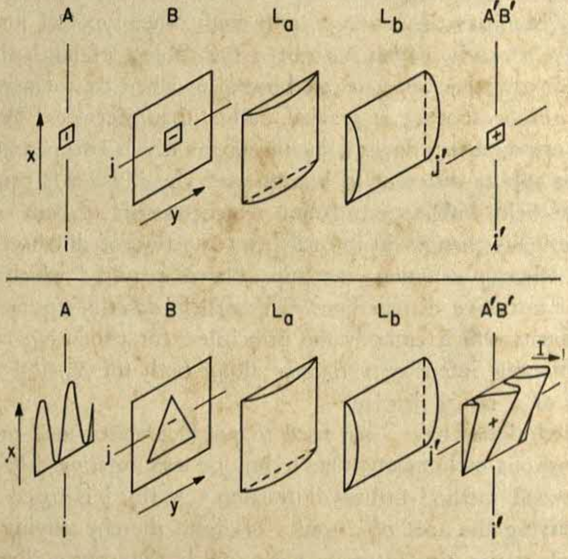


FIG. 1. SIMPLIFIED OPTICAL SYSTEMS WHICH ILLUSTRATE THE PRINCIPLES HERE DESCRIBED

image is determined by the height of the source and the width of the aperture. In general,

$$I_{i'j'} = k\, x_i\, y_j \qquad [1]$$

where $I_{i'j'}$ is the intensity of the image at point $i'j'$, x_i is the height of the source at the vertical i, and y_j is the width of the aperture at the horizontal j.

If the source and aperture in the top of Fig. 1 are considered to be single elements of a larger source and a larger aperture, it may be seen that a variety of light distributions can be built up by manipulating the appropriate dimensions of the different source and aperture elements. For example, consider the bottom of Fig. 1. The height of the source varies sinusoidally as a function of the horizontal position, i. The width of the aperture varies linearly as a function of the vertical position, j.

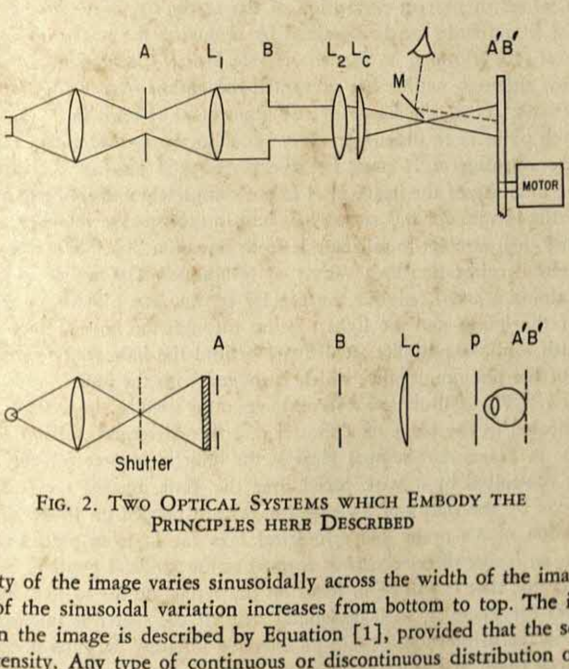


FIG. 2. TWO OPTICAL SYSTEMS WHICH EMBODY THE PRINCIPLES HERE DESCRIBED

The intensity of the image varies sinusoidally across the width of the image and the amplitude of the sinusoidal variation increases from bottom to top. The intensity at any point in the image is described by Equation [1], provided that the source is of uniform intensity. Any type of continuous or discontinuous distribution of intensity can be produced by making x_i and y_j vary appropriately.

If a filter is placed in the pathway, the intensity at every point in the image is multiplied by the transmittance of the filter. In other words, the value of k in Equation [1] is proportional to the transmittance of the filter.

Two conditions must be satisfied if the system is to operate as described. First, each point in the source must illuminate every point in the aperture equally. In other words, the aperture must be uniformly illuminated regardless of the shape of the source. Secondly, all of the light passing through the aperture must pass through the lenses. To the extent that light from a particular part of the aperture does not pass through the lenses, that part of the aperture is ineffective.

Figure 2 describes two instruments, used by the author, which embody the principles which have been described. Planes *A, B,* and *A′B′* are the counterparts of the critical planes in Fig. 1. In the apparatus shown in the top of Fig. 2, the source is a lamp with a ribbon filament. An image of the filament is formed in plane *A*, where the effective shape of the source is controlled by masking unwanted portions of the image. A shutter can be inserted adjacent to plane *A*. Images are formed on the horizontal axis by spherical lenses L_1 and L_2; images are formed on the vertical axis by the combination of the spherical lenses and cylindrical lens L_c. Combining spherical and cylindrical lenses, rather than using cylindrical lenses only, takes advantage of the higher optical quality generally available in spherical lenses. The size of the image at *A′B′* depends upon the choice and placement of the lenses. In the author's apparatus, the maximal size of the image is 1.5 by 4.0 cm. The image is viewed by reflection in mirror *M* from a distance of 25 cm. The screen is mounted on a motor-driven color wheel to prevent perception of the screen's texture. A larger field and higher retinal illuminance can be obtained by removing the screen and viewing the aerial image at *A′B′* through an eye-lens. There are two limitations associated with this apparatus, although neither has prevented the author from using the instrument in several studies.[8] First, the image of the filament at *A* must be fairly small (*e.g.* 1 cm.) if each point is to illuminate every point in the aperture at *B* equally. Consequently, the opening at *A* must be constructed and positioned with precision. Secondly, the intensity of the image at *A* is not completely uniform, particularly near the edges of the image. For this reason it is best to measure the intensive distribution at *A′B′*, rather than attempt to calculate it from Equation [1]. There is enough light at *A′B′* to permit calibration by a variety of techniques. The author has substituted a plate containing a small hole for the regular screen. The plate is moved about in the plane of the image and the light passing through the hole at each position is measured with a photomultiplier. A diffuser behind the hole scatters the light over the surface of the photomultiplier, which is moved with the hole.

The bottom of Fig. 2 illustrates a second apparatus used by the author. Light from a lamp is focused in the plane of a shutter and then diverged to illuminate a sheet of opal glass in Plane *A*. The opal glass is the effective source for the system and its shape is controlled by a mask paced over the glass. Images are formed on the horizontal axis by the optics of *S*'s eye. Images are formed on the vertical axis by the combination of *S*'s optics and cylindrical lens L_c. L_c is so placed that, on the vertical axis, an image of the source is formed at the artificial pupil *p*. The aperture at *B* is so located that it will be in focus on the retina, on the vertical axis, when *S* accommodates for Plane *A*. Each horizontal strip of the source is imaged as a horizontal strip at the artificial pupil and its effective intensity depends upon the width of the artificial pupil at the image. Because different horizontal strips of the source are imaged at different places in the artificial pupil, the effective intensity will vary from strip to strip. The effective intensities can be calibrated physically by placing a suitable sensor of light at the artificial pupil and exposing one part of the source at a time. These measurements do not allow for the Stiles-Crawford effect and for precise work they should be supplemented by individual visual measurements in which *S* matches each portion of the source with a standard stimulus for brightness.

---

[8] J. P. Thomas and C. W. Kovar, The effect of contour sharpness on perceived brightness, *Vision Res.*, 5, 1965, 559-564; J. P. Thomas, Threshold measurements of Mach bands, *J. opt. Soc. Amer.*, 55, 1965, 521-524; Brightness variations in stimuli with ramp-like contours, *J. opt. Soc. Amer.*, 56, 1966, 238-242.

Because the images on the horizontal and vertical axes are formed separately, it is possible to magnify one dimension of the stimulus more than the other. This ability is particularly advantageous when it is desired to elongate one dimension of the stimulus, but to retain along the other dimension the precision and ease in controlling size which is associated with low magnification.

APPARATUS NOTE

OPTICAL CALIBRATION OF THE FRONTAL PLANE

In a number of experimental and practical situations, the precise calibration of the physical fronto-parallel plane is required for proper control of visual direction of stimulus-targets. This note describes a simple optical method of calibration which eliminates some of the expensive engineering often required by other methods.

The apparatus is diagrammed in Fig. 1*A*. A reticle (*R*) is accurately drawn on an optically flat reflecting surface (*M*) seated in the frame of the apparatus which holds the field to be positioned in the frontal plane.[1] *O* views the mirror through an artificial pupil and small transilluminated reticle (*P*). The artificial pupil is illuminated by source (*S*) whose light is diffused by the diffusing screen (*F*) and reflected from the interface of the beam-splitter (*B*). The eye sees a projected image (*I*) on the reflecting surface (*M*) which is almost exactly one-half the linear size of the artificial pupil.

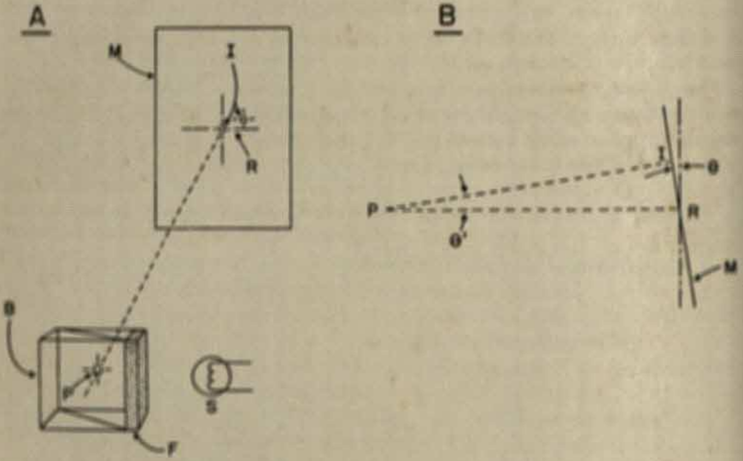


FIG. 1. GEOMETRIC RELATIONSHIP IN THE CALIBRATION OF THE FRONTAL PLANE

The mirror is positioned horizontally and vertically perpendicular to the primary visual axis by adjusting the rotation of the mirror about its two principal axes until the image of the pupillary reticle is aligned with the target reticle on the mirror.[3] The sensitivity of adjustment of the mirror-plane is dependent solely upon the visual resolution of the vernier pattern, as shown in Fig. 1*B*. The slant θ of the mirror from the true fronto-parallel plane is equal to the deviation angle θ' formed by the primary visual axis (*PR*) and the visual direction (*PI*) of the reflected image. The reliability of the settings ($1/\theta_o$) therefore depends upon the reliability of judgments of the visual direction ($1/\theta_o'$) of the reflected image (*i.e.* visual acuity).

---

[1] The reticle also may be transilluminated from the rear, with the reticle-aperture at the second surface of the glass. This procedure has the advantage that the rest of the field may be darkened, facilitating the discrimination of the reflected artificial pupil.

When these calibrations are completed, the radial rotation of the mirror (about the primary visual axis) can be calibrated by rotating the mirror about either its own horizontal or vertical axis as far as possible without losing image (I). If the mirror is not radially aligned, the image will be displaced from the adjacent reticle-ine. The sensitivity of the setting depends solely upon the slant (θ) of the mirror and the sensitivity ($1/\alpha_o$) of the eye. Threshold rotation (p_o) of the mirror about PR is given by the equation

$$\varrho_o = \sin^{-1} (\sin \alpha_o \cot \theta).$$

It is recommended that the accuracy of calibration of the mirror be read for several slants about its horizontal axis.

This method of calibration may be used for a number of applications requiring accuracy to within a few minutes of arc. These may include investigations of shape and slant, vernier acuity, cyclotorsion, figural after-effects, and visual illusions.

Pennsylvania State University ROBERT B. FREEMAN, JR.

---

[3] In Fig. 1*A*, the image is shown as it appears when the mirror is aligned, and also as it appears when the mirror is slightly rotated clockwise about its vertical axis and forward about its horizontal axis.

NOTES AND DISCUSSIONS

COMMENTS ON HELSON AND MASTERS' "STUDY OF INFLECTION-POINTS IN THE LOCUS OF ADAPTATION-LEVELS AS A FUNCTION OF ANCHOR-STIMULI"

In a paper recently published by Helson and Masters, a lifted weights experiment was reported concerning shifts in adaptation-level (*AL*) as a function of varying anchor-size from 1.56 to 2879 gm. with constant stimulus-series of 100, 150, 200, 250, and 300 gm.[1] The investigation of the empirical anchor-*AL* relationship is of high theoretical relevance in respect to Helson's well-known *weighted log mean model of AL* founded twenty years ago.[2] For Helson's mathematical prediction from his basic *AL*-formula in the year 1947 (repeated 1964) that "with constant series and variable *C* (*i.e.* size of anchor), the *AL*-varies with the fourth root of *C*,"[3] represents an exact quantitative statement explicitely derived from the formula for lifted-weights which should be empirically tested in order to examine the validity of the quantitative *AL*-model. Interestingly enough, before Helson and Masters' recent publication this very mathematical implication has never been specifically tested and verified by means of modern systematic, *i.e.* trend-analytical experimentation. So far, there have been reported only incidental experimental findings; *e.g.* summarized by Guilford[4] and mentioned, partly, by Helson himself,[5] which, however, indicated some evidence *against* this clear prediction. In fact, in their study (being the *first* trend-analytical investigation undertaken by Helson's school as to the weighted average *AL*-model from 1947), Helson and Masters have found two *inflection-points* in the locus of *AL* with relatively *extreme* anchor size; namely, at ca. 25 and 1600 gm.[6]—an empirical finding which is not consistent with Helson's statement cited above.

---

[1] Harry Helson and H. G. Masters, A study of inflection-points in the locus of adaptation-levels as a function of anchor-stimuli, this JOURNAL, 79, 1966, 400-408.

[2] Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

[3] Helson, *op. cit.,* 5; also cited in Helson, *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior,* 1964, 131.

[4] J. P. Guilford, *Psychometric Methods,* (2nd ed.), 1954, 313.

[5] Helson, *opp. citt.,* 1947, 4 and 12 f.; 1964, 129 and 139-142.

[6] Helson and Masters, *op. cit.,* 403.

Based upon an experimental study undertaken independently by Helson and Masters, exactly the same finding of two inflection-points on the empirical anchor-*AL* trend for lifted-weights has been reported by the present writer.[7] In both studies it has been found that this *cubic* anchor-*AL* relationship is directly connected with the further important result that Helson's *"constant"* weighting coefficient[8] as to anchor and series effectiveness on *AL* is, actually, a variable with changing size of anchor.[9] In spite of their conforming results, opposite conclusions have, however, been drawn by different writers in respect to the empirical validity of the quantitative *AL*-model. In the spirit of a prolific scientific development of quantitative *AL*-theory the following critical remarks on Helson and Masters' interpretation of their data may be made.

(1) Mathematical implications. Mathematically, inflection-points are, *on principle,* not predictable from Helson's quantitative model, a fact which is conceded by Helson and Masters at the beginning of their paper. They write:

> We might expect *AL* to increase monotonically with decrease in intensity of anchors above series stimuli and to decrease monotonically with decrease in intensity of anchors below series stimuli in accordance with Equation [1]:
>
> $$A = k(S^{p}B^{q}R^{r}) \qquad [1]$$
>
> where A is the adaptation-level, k is a constant, S the log mean of the series stimuli, B is the anchor stimulus, R refers to residual stimuli, and p, q, and r are weighting factors, the sum of which is taken to be 1.0.[10]

Nevertheless, at the article's end, the authors state:

> that it was predicted (*sic*!) from the weighted log-mean definition that an inflection-point would be found with anchors below the series stimuli, and this deduction was verified experimentally. . . .[11] The presence of inflection-points in the locus of *AL*s does not invalidate the weighted log-mean definition.[12]

These contrary statements by the authors, at the beginning and at the end of their paper, have become possible in consequence of misleading confusion of *mathematical*

---

[7] "8. Tagung für experimentell arbeitende Psychologen" (8th Meeting for experimentally working psychologists), Munich, April 5, 1966; see Viktor Sarris, Die Abhängigkeit des Adaptationsniveaus von Ankerreizen, *Z. exp. angew. Psychol.* 14, 1967, 1-53; also Sarris, Adaptation-level theory: Two critical experiments on Helson's weighted average model, this JOURNAL, 80, 1967, (in press).

[8] See Helson, *op. cit.,* 1947, 3 ff.; 1964, 129-133; also Helson, Adaptation-level theory, in Sigmund Koch (ed.), ***Psychology: A Study of a Science: 1, Sensory, Perceptual, and Physiological Foundations,*** 1959, 579-584.

[9] Helson and Masters, *op. cit.,* 405-407; Sarris, *op. cit.,* 1967, (in press).

[10] Helson and Masters, *op. cit.,* 400. For sake of clarity it is stressed here that *up to this point,* at least, Equation [1] fully corresponds to Helson's earlier expositions about the weighted log mean definition of *AL* (see Footnote 8), even though at this point the authors only implicitly assume the weighting factors as to anchor-*AL* relationship to be *constants.*

[11] Helson and Masters, *op. cit.,* 407.

[12] Helson and Masters, *op. cit.,* 408.

derivation and *psychological reasoning.* This objection to Helson and Masters' interpretation of their data is confirmed by the following analysis.

In contrast to Helson's basic writings on the weighted log-mean AL-model,[13] Helson and Masters do not specify the weighting coefficients p, q, and r as *constants* (they do not mention Helson's earlier publications in respect to this point), but they tacitly omit this problem until the experimental results are discussed; and it is simply stated, then, "that the weighting coefficients, p, q, and r . . . are not constant over the entire range of stimulus-intensities."[14] This omission has been made *up to this point,* excepting perhaps the special case of $B = O$, where the authors argue, not by means of mathematical derivation from Equation [1], but by means of psychological reasoning (empirical plausibility). They state: "It is clear that if there is no background or anchor-stimulus, the term B^q reduces to 1.0 because a zero stimulus is weighted *zero* and $0^\circ = 1.0$."[15] Helson and Masters, at this point in their paper, do *not* assume (or "derive"?) explicitly a *continuum* within which the weighting factors may vary, *e.g.* between 0 and 3.0 or so;[16] they merely suggest, again from psychological reasoning, that there must be "a certain point after which *AL* must reverse and approach series *AL* as the anchor approaches zero."[17] Tacitly assuming that a "no anchor-condition" is psychologically identical with "zero anchor condition" as to *S*s judgments. On the other side, their investigation, later remarked in their *Discussion,* —"that the weighting factors, p, q, and r, may vary within a given dimension, especially at the extremes of the stimulus-continuum,"—[18] is, at least, not consistent with Helson's earlier assumption that loss of anchor effectiveness somewhere between 5.0 and 0.5 gm. "introduces a *discontinuity* in the weighted log mean definition of *AL*".[19] It is difficult to understand fully Helson and Masters' whole line of reasoning;[20] but, all in all, mathematically arguing, the inflection-point at the lower end of the stimulus-continuum has *not* been "predicted" from Equation [1].

At this point, it may be summarized that Helson and Masters do not adequately relate their experimental findings; (a) to their Equation [1]; and (b) to Helson's earlier statements cited above. Corresponding to this, it appears as an important shortcoming of Helson and Masters' work that the authors have not precisely measured and explicitly used the *variable* weighting coefficient as to series and anchor effectiveness in their Equation [1], though they state: "We thus have a quantitative measure of relevance in the values of the weighting coefficient for series, anchor, and residual stimuli."[21] This means, the authors are satisfied with a *global,* qualitative declaration of p, q, and r being variable in character. They do not specify the variable weighting coefficients (as to anchor and series) as an asymmetric *parabola* dependent on the anchor value as the present writer has done elsewhere.[22] They do not mention any feasible modification of the quantitative AL-formula from 1947. But this failure necessarily excludes the possible mathematical *"prediction"* of inflection-points as well at the lower as at the upper anchor-range. As it is implicitly demonstrated in Helson and Masters' paper, however, unless the basic mathematical structure of the respecting weighting variable has been taken into appropriate account, it is not *AL*-theory which is invalidated but the *quantitative* model; namely, Equation [1].[23]

---

[13] See Footnote 8.

[14] Helson and Masters, *op. cit.,* 405.

[15] Helson and Masters, *op. cit.,* 400. (Italics ours.)

[16] It should be noted to the reader, perhaps, that Helson's empirically established *constant* weighting factor as to anchor-*AL* relationship for lifted weights has been $k = 3.0$ (see Helson, *op. cit.,* 1947, 4; also 1964, 130). Throughout their paper Helson and Masters do *not* refer to this constant $k = 3.0$.

[17] Helson and Masters, *op. cit.,* 400.

[18] Helson and Masters, *op. cit.,* 405.

[19] Helson, *op. cit.,* 1964, 142. (Italics ours.)

[20] For example: What does Helson and Masters' tacit omission of the "*d*-factor" throughout their paper actually mean? If the authors would have wanted to make an *exact* test of Helson's weighted log-mean model as to anchor-AL relationship, they necessarily should have taken into account this "correctional" factor (see Helson, *op. cit.,* 1947, 4-7 and 9-11; also *op. cit.,* 1964, 130-135). Or does the omission imply a tacit *fundamental* rejection of this "arbitrary" factor recently criticized, by Parducci (Allen Parducci, Adaptation-level theory, this JOURNAL, 78, 1965, 158-159)?

Why do Helson and Masters not revise the *AL*-model *quantitatively* in respect to the weighting factors? Is it that Helson dislikes a *complex* mathematical model? Fundamentally asking, why is it that Helson over years has been content above all with *simplicity* of a quantitative model?[24] For he has argued, *e.g.*" "A simple mathematical formulation may have more power than a complicated mathematical model if the former deals with fundamental mechanism of behavior and the latter is concerned with minutiae or trivia."[25] Though this statement has its own rights, one should not forget, nevertheless, that also a simple mathematical formulation must be *correct,* because, as Stevens has put it in his argument against *AL*-theory in another connection, "a quantitative theory that is wrong is not better for being quantitative."[26]

(2) Significance of inflection-points. As stated above, Helson and Masters have "predicted" the inflection-point at the lower end of the scale by means of a mathematical "trick." The inflection-point at the high end of the scale is explained by means of the assumption," that different muscles and a different stance are employed in lifting the anchor-stimuli with resultant change in judgment" and, besides, "that the lifting situation is considerably different for the anchor and series stimuli."[27]

This *a posteriori* hypothesis seems to have two disadvantages: (a) It gives an explanation, if any, only for lifted-weights, but not for other psychophysical materials where inflection-points may be found as well. It means, this hypothesis lacks *generalizing* power. For example: how should inflection-points at the high end of the scale for *pitch* be interpreted?[28] (b) This hypothesis is not convincing even for lifted-weights themselves. The present writer found in his own lifted-weights experiment, with series stimuli of 200, 250, 300, 350, and 400 gm., two inflection-points with anchor of ca. 75 and 2900 gm.;[29] *i.e.* both inflection-points had shifted *upwards* in comparison with Helson and Masters' inflection-points at ca. 25 and 1600 gm. (see above). It appears rather unlikely that Helson and Masters' "different muscles" are, with changing stimulus-series, for one time employed at 1600 gm., and for the other time not before 2900 gm. But, of course, it is difficult to compare the experiments of Helson and Masters with those of the present writer upon a common reliable base. Though Helson's hypothesis seems not to be refutable at the present scientific stage, the writer's 'class' hypothesis,[30] according to Helson's own line of psychological reasoning,[31] seems to be more likely, because it gives more sense for an experiment with *pitches* where variation of both series and anchor produced a clear trend of *shifting* inflection-points varying with size of stimulus-series.[32]

---

[21] Helson and Masters, *op. cit.,* 407.

[22] By means of using the empirical *parabolic* relationship between anchor size and weighting variable coefficient for anchor and series in Helson's basic *AL* formula the inflection-points become "predictable." Moreover, as the present writer has shown, this *modification* has important theoretical advantages for the quantitative *AL* model, as well; see Sarris, *op. cit.,* 1967.

[23] This point has been treated by the present writer in detail elsewhere; see Sarris, *Z. exp. angew.* Psychol., 14, 1967, 43-44; also this JOURNAL, (in press).

[24] See *e.g.* Helson, *op. cit.,* 1947, 27 f.; *op. cit.,* 1959, 576; *op. cit.,* 1964, 24-26.

[25] Helson, *op. cit.,* 1959, 576.

[26] S. S. Stevens, Adaptation-level vs. the relativity of judgment, this JOURNAL, 71, 1958, 634.

[27] Helson and Masters, *op. cit.,* 406.

(3) The experimental procedure. It is interesting to note that Helson and Masters have actually *not* proved both "inflection-points" to be statistically significant phenomena.[33] There is only "graphical" evidence for inflection-points and, nevertheless, Helson and Masters have interpreted them as real facts. Instead of using successive *t*-tests the authors better should have used *trend analysis.*[34] Since in the writer's study inflection-points were proved to be statistically relevant, Helson and Masters' questionable interpretation of their data may be accepted, nevertheless.[35] Apart from this, Helson and Masters' experimental procedure is, perhaps, not the most appropriate one because of the following points:

(*a*) Though "carry-over effects" must be supposed for the lifted-weights situation, the authors nevertheless have used a so-called "repeated measurements design," *i.e.* a design providing repeated measurements for the *same* individuals. Thus Helson and Masters have had their *S*s (11 in number) judging the stimulus-series under all experimental conditions, the anchors being introduced to *S*s in randomized sequence.[36] Though carry-over effects may have been of little influence in comparison with treatment (anchor) effects, an expert like Winer has generally argued *against* this type of design as follows: "A strong word of warning is required in connection with order (or sequence) effects. . . . If such effects exist, randomizing or counterbalancing does not remove them; . . . in experiments . . . where the sequence effects are likely to be marked and where *primary interest lies in evaluating the effect of individual treatments in the absence of possible sequence effects,* a repeated-measures design is to be avoided."[37]

---

[28] Recent systematic experiments with *pitches* by the present writer concerning shift in *AL* with variation of both anchor *and* series support the presence of inflection-points with anchor becoming extreme in comparison to stimulus-series most clearly at *both* ends of the scale. The report of this research work is in preparation.

[29] Sarris, *opp. citt.,* 1967.

[30] Sarris, *op. cit., Z. exp. angew. Psychol.,* 1967, 45; *op. cit.,* (in press).

[31] Helson, *op. cit.,* 1964, 126 f.

[32] See Footnote 28.

[33] Helson and Masters, *op. cit.,* 404.

[34] As to the methodological shortcomings of successive t-testing see *e.g.* Guilford, *Fundamental Statistics in Psychology and Education,* (3rd ed.), 1956, 257 f. As to the power of *trendanalysis* see E.G. Quinn McNemar, *Psychological Statistics,* (3rd ed.), 1962, 350 f.

[35] Sarris, *op. cit., Z. exp. angew. Psychol.,* 1967, (in press).

[36] Helson and Masters, *op. cit.,* 402.

(*b*) In the experimental session, Helson and Masters' *S*s have been instructed, among others: "Later you will be given other series of stimuli to judge that will be heavier or lighter than the present series."[38] In expecting *heavier* or *lighter* series for subsequent experimental phases, the *S*s may have been influenced by this remark which calls attention to more extreme stimuli. As Helson himself has suggested elsewhere, *verbal* instructions may alter the conditions of a "stimulus field,"[39] and Guilford states: "An imagined stimulus may also serve as an anchor."[40]

(*c*) In the control session ("no anchor"), *E* presented two practice runs to the *S*s, in the subsequent anchor session only one practice run.[41] It may well have been that the *S*s were not experienced enough for a special experimental condition to give reliable judgments; for, as Rambo and Johnson have shown, there is an important dependence of *AL*-values from practice.[42]

Reviewing Helson and Masters' article, one is inclined to assume that there investigation has not overcome Parducci's general critical comment against *AL*-theory recently made: "Considered as a quantitative theory, its evaluation must await further development and testing."[43]

University of Duesseldorf VIKTOR SARRIS

STIMULUS-VARIATION AND RECALL: THE ROLE OF BELONGINGNESS

The authors of this Note have presented evidence that some variation in the material provided for learning with repeated practice is a more favorable condition for the recall of a generic stimulus than is exact repetition of the stimulus in all its detail.[1] Whle repetition is a principal instrument in the acquisition of particular responses, variation becomes the

---

[37] J. B. Winer, *Statistical Principles in Experimental Design,* 1962, 301. (Italics ours).

[38] Helson and Masters, *op. cit.,* 402.

[39] Helson, *op. cit.,* 1964, 327.

[40] Guilford, *op. cit.,* 1954, 313.

[41] Helson and Masters, *op. cit.,* 402.

[42] W. W. Rambo and E. L. Johnson, Practice effects and the estimation of adaptation-level, this JOURNAL, 77, 1964, 106-110.

[43] Parducci, *op. cit.,* 159.

* Partly supported by Contract Nonr-3634(01) between the Office of Naval Research and Kansas State University. The experiment was carried out while the senior author was in residence as a National Science Foundation Senior Postdoctoral Fellow at the Center for Advanced Study in the Behavioral Sciences.

[1] William Bevan, W. F. Dukes, and L. L. Avant, The effect of variation in specific stimuli on memory for their superordinates, this JOURNAL, 62, 1966, 250-257; W. F. Dukes and William Bevan, Stimulus variation and repetition in the acquisition of naming responses, *J. exp. Psychol.,* 1967 (in press).

key to acquiring concepts. But the proper heuristic consideration is not to ask whether variation is important but what kind of variation is required. Some of the early work on contextual stimuli and recall suggests that the nature of the influence of stimulus change will be contingent upon what is varied.[2] The present paper focuses on the role of one particular type of variation; namely, variation is the usualness of the stimulus-context as a condition of recall. Taking as a starting point Thorndike's concept of belongingness, one might hypothesize that if x is presented once with a and once with b, the variation will be more beneficial in recalling x if a and b "belong with" x than if they do not.[3] The experiment reported here tests that notion.

The *Ss* were members of undergraduate psychology and education classes at the University of California, Davis. Because the proportion of men to women varied widely in the classes tested and because the scores of the men, although consistently lower, manifest the same trend as those of the women, the data analysis presented below and its attendant discussion are based only on the records of the latter ($N =$ 261).

The experiment was conducted in regular classes ranging in size from 9 to over 100. The only instructions given the *Ss* were that pairs of common English words would be projected briefly on the screen before them. Forty slides were then shown each group, with a 5-sec. exposure-time per slide, each slide carrying an adjective-noun pair. Order of presentation was random for the first 20 slides (each noun presented once); the last 20 slides were so arranged that no noun was in a highly favored or unfavored position in both presentations, and no noun either succeeded or preceded the same noun as it did in the first 20 slides. The order of nouns was identical for each experimental condition.

After the words had been presented, each *S* was given a sheet of paper and asked to write down as many of the nouns as he could remember; 1.5 min. were allowed for recall. When these papers had been collected, the procedure was repeated for the adjectives.

Stimuli were 20 nouns and 40 adjectives.[4] All of the nouns appeared twice in every experimental condition. Each adjective was paired once with a noun to which it 'belonged' and once with a noun to which it did not belong, but no adjective was used with more than one noun under any one experimental condition. Thus, under some conditions (see *Design*) both adjective and noun—the stimulus pair—were repeated, only one-half of the adjectives being used, while under others, only the nouns were repeated, all 40 adjectives being used. Whether an adjective belonged with a noun was determined by *E*'s judgment. Common or expected combinations—*e.g.* green tree, cold wind—were regarded as belonging; unusual combinations—*e.g.*

---

[2] Shuh Pan, The influence of context upon learning and recall, *J. exp. Psychol.*, 9, 1926, 468-491; Helen J. Reed, The influence of a change of conditions upon the amount recalled, *J. exp. Psychol.*, 14, 1931, 632-649.

[3] E. L. Thorndike, *The Fundamentals of Learning*, 1932, 64-72.

[4] The nouns were selected from the 500 most frequently occurring words in E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 words*, 1944, and 31 of the adjectives from the first 1,000. The remaining nine were common adjectives selected by the *Es* to permit use of certain noun-adjective combinations.

low paper, sudden school—although perhaps reasonable in some contexts, were placed in the *not belonging* category. The paired words were on slides for use in a Kodak Carousel 35-mm. projector.

Design. A basic 2 × 2 design was used involving the following four conditions. Condition (1) *Belonging* adjective-noun pairs *repeated—e.g.* "tall tree" occurs in the first 20 slides and is repeated in the last 20. Condition (2) *Belonging* adjective-noun pairs *varied—e.g.* "tall tree" becomes "green tree" on the second presentation. Condition (3) *Non-belonging* adjective-noun pairs *repeated—e.g.* "sweet army" occurs in the first 20 slides and is repeated in the last 20 and Condition (4) *Non-belonging* adjective-noun pairs *varied—e.g.* "sweet army" becomes "round army" on the second presentation. Conditions (1) and (3) each involved two sub-groups of *S*s, since in repeated adjective-noun pairings only one-half the adjectives would be used in any one experimental group. Sub-groups could also have been used in Conditions (2) and (4) by varying the order of the pairing, *i.e.* tall-green, green-tall—but this was deemed of no consequence for our present purpose.

TABLE I

MEAN RECALL SCORES

Conditions	*N*	Nouns	Adjectives
(1) Belonging pairs repeated	35	10.06	10.06
(2) Belonging pairs varied	65	12.35	14.97
(3) Non-belonging pairs repeated	99	10.45	9.19
(4) Non-belonging pairs varied	62	11.16	10.69

Recall scores—one for nouns, another for adjectives—were determined for each *S* by assigning one point for each item recalled. Thus for nouns a score of 20 was possible and for adjectives 20 under repeated conditions, 40 under varied. Mean scores obtained under the various conditions are given in Table I. The sub-groups under the conditions of repeated presentation—Conditions (1) and (3), above—were combined, there being no significant difference between them. An analysis of variance was performed on the data for nouns, the crucial material for the hypothesis under test. Three principal findings are evident:

(1) Variation in the adjectival modifiers enhances recall of the nouns. Condition (2) exceeds Condition (1), and Condition (4) exceeds Condition (3) ($F_{variation}$ (1257) = 20.45, $p < 0.01$). This finding confirms previous work of the present authors on memory for generic stimuli.[3] The repetition of the adjectives does have an effect on their recall. About 50% of the adjectives were recalled under the *repeated* conditions (10.06 and 9.19 out of 20), but only about 36% of them were recalled under *varied* conditions (14.97 and 10.69 out of 40).

(2) Belongingness, as defined here, did not affect recall in a simple, unqualified way. Condition (2) exceeds Condition (4); although Condition (3) does not differ noticeably from Condition (1) ($F_{belonging}$ (1257) = 1.45 $p, > 0.05$). This finding is at first a little surprising. There

---

[3] Bevan, Dukes, and Avant, *op. cit.*, 252-253.

may have been 'spill-over' among pair members, encouraged by the design. That is, the same modifiers are used both in the *belonging* and *not-belonging* conditions, such that *a,* though paired with *x,* may facilitate the recall of *y* because of links established prior to the experiment. The superficial contradiction is further resolved by considering the interaction described below.

(3) Clearly the type of variation is important. Consonant with the initial hypothesis, variation is a more effective agent in recall when the adjective varied 'belongs with' the noun to be recalled than when not ($F_{\text{interaction}}$ (1257) $= 5.64$, $p < 0.05$), note the 2.29 advantage of variation under *belonging* conditions vs. the 0.71 advantage under *non-belonging.*

It is tempting to regard this finding as favoring an associationistic hypothesis over a motivational one: that is, that variation is effective more because of a greater number of links than because of the interest provided by diversity. Thus, the varied adjectives provide potentially twice the number of possible associations, and the adjectives that belong to the nouns presumably represent stronger associative bonds. However, partly because of the possibility of 'spill-over' mentioned above and partly because of the supplementary data on adjectives, one can not wholeheartedly accept this interpretation. The *belonging-repeat* group recalled 10.06 adjectives and 10.06 nouns while the non-belonging-repeat group recalled 9.19 adjectives and 10.45 nouns. Finally, it is important to note that under the repeated conditions, belongingness had no identifiable influence on recall (10.06 vs. 10.45, t_d (132) $= 0.81$, $p > 0.05$).

SUMMARY

A total of 261 women, tested in groups, were presented briefly with noun-adjective pairs. Following these presentations, they were first asked to list the nouns they could remember and later the adjectives. Two groups of *S*s saw the same pairing of adjective with noun in the second half of the series; two other groups saw different combinations on the second presentation. As a second variable, the belongingness of the adjective to the noun was altered. Variation proved superior to repetition as a condition of recall of the nouns regardless of whether the adjective belonged to the noun or not. Variation was most effective when the pairs were high in belongingness. In contrast, belongingness exerted no influence upon recall of the nouns under the repeated condition. Recall of the adjectives was better when the combination of the adjective with a noun was repeated.

The Johns Hopkins University WILLIAM BEVAN
University of California, Davis W. F. DUKES

ALTERNATIVE INTERPRETATIONS OF CUTANEOUS PERCEPTION

Since Krech and Crutchfield's introduction of the method of finger-tracings on the head and their interpretation of the results in terms of the locus of the "self" (external versus internal)[1], the method has been used successfully in several investigations.[2] Consistent with the view of Krech and Crutchfield, these experiments have been interpreted using a theoretical concept of the perceiver (ego) whose properties include an ability to change locus and orientation.

Recently Yonge offered an alternative interpretation.[3] He also criticized the conclusions of Natsoulas and Dubanoski as unparsimonious and as inapplicable to the results of an experiment by Dubanoski involving tracings on the tops of subjects' heads.[4] Yonge's critique has stimulated the application of the original interpretation to the latter results; a re-analysis of the data provided clear support.[5]

As far as a choice between the alternative interpretations is concerned, it will depend on which one best fits the results of empirical investigations. Thus far, Yonge's interpretation has not been applied as widely as has Krech and Crutchfield's as developed by Natsoulas and Dubanoski. This is not to say it cannot be applied; only that it has not been.

Certain difficulties moreover appear for Yonge's schema. First of all, there appears to be greater consistency in response to tracings on the sides of the head (with head facing straight ahead) than Yonge's concept of "confusion" of figure with its mirror-image would allow in its present unqualified form.[6]

Secondly, and of greater importance, is the inability of Yonge's schema to meet an empirical test which he himself has proposed. The test involves having *S* point to the left (with eyes closed and head turned to the right). Yonge predicted and appears to have found that "*S* points somewhere *between* the normal forward and far-left axis. Corresponding results are obtained when *S* is asked to turn his head to the left."[7] Yonge gives no additional details of the procedure or results. An effort to verify these claims was not successful.[8] No one pointed, as Yonge expects, *partway* between the frontal and lateral directions.

---

[1] David Krech and R. S. Crutchfield, *Elements of Psychology,* 1958, 205. See also F. A. Geldard's comments (*Fundamentals of Psychology,* 1962, 279) on the use of "figure writing" in clinical neurology as "part of a battery of tests of integrity of the nervous system."

[2] Thomas Natsoulas and R. A. Dubanoski, Inferring the locus and orientation of the perceiver from responses to stimulation of the skin, this JOURNAL, 77, 1964, 281-285; Thomas Natsoulas, Locus and orientation of the perceiver (ego) under variable, constant, and no perspective instructions, *J. person. soc. Psychol.,* 3, 1966, 190-196; Thomas Natsoulas and J. T. Murphy, Locus and orientation of the perceiver (ego) and the rotation of visual images, *ibid.,* 4, 1966, 472-475.

[3] G. D. Yonge, A reinterpretation of the Natsoulas and Dubanoski study, this JOURNAL, 78, 1965, 677-683.

[4] R. A. Dubanoski, Inferring the phenomenal locus of the peceiver from responses to tactile stimulation, Unpublished Master's thesis, University of Minnesota, 1964.

[5] Thomas Natsoulas, On the perception of cutaneous figures, *Psychol. Rec.,* 1967 (in press.)

[6] *Ibid.*

Perhaps there are special instructions, justified by Yonge's schema, which will give the required results. Perhaps a more subtle method would show that the bodily left-right axis turns when the head turns.

University of California, Davis THOMAS NATSOULAS

Kenneth Wartinbee Spence: 1907–1967

In the death of Kenneth W. Spence on January 12, 1967, a few months prior to his sixtieth birthday, psychology has suffered the loss of one of the leading experimentalists and theorists of his generation. Having moved in 1964 from the State University of Iowa to the University of Texas, he was busily involved in research in his new laboratory, in writing and editing, close to the time of his death by cancer.

Kenneth was born on May 6, 1907 in Chicago but moved to Montreal at the age of four or five years, where his father, an electrical engineer, worked until retirement. He was very active in sports as a boy and young man, a basketball star in high school and a track star in college. He won trophies in tennis and bowling. During his freshman year at McGill, he injured his back in track activities; as part of his convalescence, he went to live with his grandmother in LaCrosse, Wisconsin, where he attended LaCrosse Teachers College (now LaCrosse State University), majoring in physical education. His bad luck followed him, however, and a fractured skull and jaw received in a push-ball contest landed him in the hospital for six weeks. It was in LaCrosse that he met and married Isabel Temte, who became the mother of his two children, Shirley Ann (now Mrs. Donald Pumroy) and William James.[1] He returned to McGill to complete his undergraduate education (B.A., 1929), and went on for a Master's degree (1930), shifting to psychology as his major, although he continued to support himself in part with jobs as a playground director and basketball coach for a YMCA team in Montreal. In the summer of 1930, before going to Yale, he worked as an interviewer in the personnel department of the Hawthorne plant of the Western Electric Company.

---

[7] Yonge, *op. cit.*, 680.

[8] Natsoulas, *op. cit.*, *Psychol. Rec.*, 1967 (in press).

[1] He and Isabel were later divorced, and in 1960 he married Janet A. Taylor.

In his graduate work at Yale, Spence became an assistant in the laboratories of Robert M. Yerkes, and completed his dissertation on visual acuity in the chimpanzee under Yerkes' direction. He continued for a year after receiving his Ph.D. in 1933 on a National Research Council Fellowship, and then went to the Yale Laboratories of Primate Biology at Orange Park, Florida, for three additional years as a research assistant with rank of instructor. While at Yale, Spence began an intellectual association with Clark L. Hull which greatly influenced his later career. As an exercise in the graduate course in experimental psychology which I was teaching at the time, he and Walter C. Shipley performed an experimental test of one of Hull's deductions concerning the maze-learning of the rat, published later in the *Journal of Comparative Psychology* (1934). This was one of several papers on rat learning published "on the side" while doing his dissertation on visual acuity in the chimpanzee. His teachers at the time recognized his great promise in designing experiments relevant to theory.

Spence was a product of the depression years, when attractive academic opportunities were scarce. His chance for a teaching career came when he was offered a one-year assistant professorship at the University of Virginia (because of someone's being on leave). That was in 1937; he had already had his Ph.D. for four years. The following year he went to the State University of Iowa, where he spent the next 26 years, 22 of them as department head, before moving to the professorship at the University of Texas in 1964.

The years at Iowa were tremendously productive, both in personal productivity and in the supervision of doctoral dissertations. Spence was responsible for no less than 73 Ph.D.'s at Iowa during the years 1940-64, an average of three a year, year in and year out.[2] While such statistics are not available for exact comparison with others, I doubt if this record can be matched in American psychology.

Spence's central and persisting interests in the experimental study of learning and in the theoretical formulation of the processes involved give him a clear and important place in the history of these topics. Not only was he greatly influenced by Clark L. Hull, but he also had a great deal of influence upon Hull's thought; in some ways the prominence of the Hullian position in the 1940s and 1950s held back the recognition of the independent contributions that Spence was making all along because he was pictured then as a disciple of Hull. Even now, the viewpoint that they jointly represented is often referred to as the Hull-Spence position. Hull acknowledged how much help he had received from Spence prior to the appearance of Hull's *Principles of Behavior* (1943) in the preface of that book, in which he said "To Kenneth W. Spence I owe a debt of gratitude which cannot adequately be indicated in this place." The most important theoretical contribution to this book was the introduction of the intervening variable, which Spence had persuaded Hull to adopt from Tolman.[3]

---

[2] With no intention of being invidious through omissions, the following 10 names have been selected as representative of Spence's students who received their Ph.D.s prior to 1950; Abram Amsel, Sidney W. Bijou, I. E. Farber, Arthur L. Irion, Howard H. Kendler, Gregory A. Kimble, Irving M. Maltzman, Janet A. Taylor, Benton J. Underwood, and Wilse B. Webb.

To describe Spence's major contributions in a short space is impossible, but a few of the highlights can serve as reminders of other things that he has also done:

(1) Discrimination learning. He early evolved a theory of discrimination learning based upon overlapping gradients of generalization, a gradient of excitatory strength from the reinforced member of the training pair and a gradient of extinction from the negative pair. The theory permitted an explanation of some transposition-phenomena and the puzzling problem of difficulty with selecting the intermediate member of a three-member stimulus-presentation, without invoking any Gestalt-like principles. Further on discrimination learning, Spence sided with the continuity theorists against the discontinuity ones, producing both experiments and theory.

(2) Mathematical formulation of motivation and discrimination. Carrying on Hull's efforts to find precision in theory through mathematics, Spence developed a number of models permitting experimental choice between alternative theoretical proposals. These are well represented in his *Behavior Theory and Conditioning* (1956), Spence's *Silliman Lectures* at Yale. His selection as a Silliman lecturer was a signal honor, as he was the only psychologist appearing in this distinguished series. He was particularly concerned with the interaction between incentive and motivational variables in animal learning, and arrived at a conception that differed from Hull's, specifically finding it desirable to *add* D (drive) and K (incentive motivation),[4] whereas Hull *multiplied* them. By manipulating these variables in a strictly quantitative manner, Spence was able to show that increasing motivational level will facilitate performance on tasks in which the correct, to-be-learned response is stronger than those of other response-tendencies elicited by a stimulus, but will deter performance on tasks in which the habit-strength of the correct response is initially weaker than those of competing response-tendencies. He showed also that the mathematical form of the curves obtained when probability of the conditioned response is plotted against successive presentations of the paired stimulus changes systematically with motivational level.

---

[3] The argument for its use was presented by Gustav Bergmann and K. W. Spence, Operationism and theory in psychology, *Psychol. Rev.*, 48, 1941, 1-14.

[4] Hull is said to have chosen *K* as the designation for incentive motivation to honor Kenneth, who brought it strongly to his attention.

While the kind of mathematics used does not have the elegance of some of the newer stochastic models, it is closely in touch with empirical data, and has a measure of flexibility in application which some of the newer models lack.

(3) Extinction in classical conditioning. Spence's preoccupation in the later years of his laboratory work was with phenomena of extinction in classical conditioning, including its integration in man with cognitive factors. It is not always recognized that various aspects of *inhibition* were Pavlov's most distinctive contributions. Throughout history, positive associations had been recognized, and, of course, forgetting, but the active role of extinction through nonreinforcement was a powerful new idea that came from Pavlov. Spence perceived that we did not yet have any very clear idea about extinction as a behavioral phenomenon, and he was engaged on a number of studies at the time of his terminal illness. Some came to publication, on cognitive factors in extinction, on number of trials and presence or absence of the unconditioned stimulus, on the discriminability of the change when extinction was begun. Doubtless this series of experiments, had they come to completion, would have been organized under a theoretical framework that would have furthered our understanding of these puzzling phenomena beyond what these studies have already contributed.

(4) Logic of science in psychology. Throughout his career, Spence was interested in what would constitute an appropriate kind of scientific logic for psychology, and he repeatedly examined the assertions of learning theorists to compare them with one another. He wrote at least a dozen articles and book chapters on these matters. He argued for a behaviorism, but recognized that there was very much in common between Tolman, Skinner, and Hull, and was not rigid about *what kind* of behaviorism was needed. He continued to argue for the economy and power of the intervening variable, having characterized such variables as follows: "intervening variables of learning theory serve as general or abstract terms that are applicable to a variety of situations."[5] He reaffirmed this position in the foreword which he prepared for the seventh printing (paperbound) of Hull's *Principles of Behavior* appearing about the time of Spence's death. In this foreword he noted how similar the earlier writings of Tolman, Skinner, and Hull were, and expressed regret that Skinner departed from his earlier intervening variable ("hypothetical middle terms") type of thinking. The dissimilarities among the writers he of course recognized, and the characterization of Tolman's as an S-S system, thus differing from Skinner's and Hull's as S-R systems, is one he originated.

Spence's early promise and later achievements did not go unrecognized.

---

[5] K. W. Spence, The empirical basis and theoretical structure of psychology, *Philos. Science*, 24, 1957, 97, 1-8.

Already at McGill he was awarded the Prince of Wales Gold Medal in Mental Sciences upon graduation in 1929, and in his first graduate year there received the Governor-General's Medal for research. He received the Howard Crosby Warren Medal awarded by the Society of Experimental Psychologists in 1953. He was elected to the National Academy of Sciences in 1955.

We are fortunate that his legacy to psychology in the form of some 100 scientific reports, book chapters, and theoretical papers, is readily accessible through two books, the Silliman Lectures, *Behavior Theory and Conditioning* (1956), and his own selection of papers that he considered most representative, up to that time, *Behavior Theory and Learning: Selected Papers* (1960).

Stanford University ERNEST R. HILGARD

AN ACKNOWLEDGMENT

The portrait and signature of Professor Kenneth W. Spence, which are reproduced in the frontispiece of this number, were kindly supplied the JOURNAL by his wife, Professor Janet (Taylor) Spence. Both facsimiles are of recent date.

K.M.D.

ERRATA

Dr. M. L. Cook, author of the study on "An exact test of the size-distance invariance hypothesis," which appeared in the December, 1966 number of this JOURNAL (Vol. 79, pp. 568-575), invites attention to and correction of the following typographical errors that were made in the formulae.

On page 571, the left-hand sides of Equations [4] and [5] should read: d_a'/d_b'; the left-hand side of Equation [6a] should read: $s_{st}'/s_{va(d)}'$

On page 572, the left-hand side of Equation [9a] should read: $\bar{d}_{va(d)}$.

On page 573, the bracketed equation in line 17 should read: $[m_{\bar{s}_{va(s)}} m_{d_{va(d)}} - covar\ (\bar{s}_{va(s)} d_{va(d)}/n)$;

the left-hand side of Equation [16] should read $\hat{D}$ instead of D, and $\hat{D}$ should appear in lines 20 and 21 instead of D;

On page 574, lines 45 and 46 should read: $m_{\bar{s}_{va(s)}}$, $m_{d_{va(d)}}$, $covar\ (\bar{s}_{va(s)} d_{va(d)})$

On page 575, in lines 4, 5, and 6, substitute $\hat{D}$ for D, Column 4 in Table I should be headed: $covar\ (\bar{s}_{va(s)} d_{va(d)})$, and Column 5 should be headed; $\hat{D}$.

Dr. Albert Silverstein, author of the study on "The prediction of individual association-heirarchies from cultural frequencies," which appeared in the March 1967 number (Vol. 80, pp. 88-94), invites attention to the error that was made in setting line 18, p. 88. The sentence beginning in the middle of that line should read: "In this method, *S* cannot *repeat* his associations to a stimulus. . . ." "Report" should be "repeat."

K.M.D.

BOOK REVIEWS

Edited by T. A. Ryan, Cornell University

An Introduction to Motivation. By JOHN W. ATKINSON. Princeton, N.J., D. Van Nostrand Co., 1964. Pp. xiv, 335. $9.75.

Motivation: Theory and Research. By CHARLES N. COFER and MORTIMER H. APPLEY. New York, John Wiley and Sons, 1964. Pp. vi, 958. $12.50.

Both of these books belong on the reading list of the advanced student in psychology—both, not one or the other, because they have different qualities to offer the reader. Atkinson's book is more tightly organized around a central theme of *Expectancy X Value* theory. Other approaches are also presented, but with the emphasis upon their contrasts and differences from the expectancy X value formulation. Cofer and Appley, on the other hand, present a wider range of points of view and more detailed research results, with a more neutral framework, which automatically entails a looser organization of the book as a whole.

Although both books represent major contributions which are mutually supplementary, each has its shortcomings. Some of these are inevitable in such an ill-defined and sprawling field as motivation and will be rectified only as time provides further thought and research. The criticisms to be mentioned are therefore not meant to reflect upon the accomplishments of the three authors, but to suggest ways in which future writers may go on to further clarity and organization.

In the first two-thirds of Atkinson's book he has set himself the important goal of summarizing accurately and fairly the thought of such forerunners as James, Freud, Lewin, Thorndike, Tolman, Hull, and their successors. In the interest of fair representation, Atkinson has quoted each of the relevant writers at considerable length, and he has treated the theoretical issues raised by these writers in some depth. Although it is very successful in its aims, it calls for close and careful study, even by a fairly sophisticated student. Undergraduates and even graduate students will find it effortful though rewarding work. This reader feels that the book could have been made more manageable for students by further selectivity, reducing the amount of quotation and detailed argumentation. Atkinson does give frequent summaries and statements of the argument that is to come, and perhaps that is the best that can be done if it is deemed essential to show the historical development of a particular idea or point of view (*e.g.* Hull's views on drive). Yet there are great advantages in presenting the basic conceptual alternatives without regard to historical development. After all, the solutions of problems will depend upon the value of the concepts, not on how they developed or their historical origin.

The last three chapters of Atkinson's book deal with recent research closely related to his own interest, his own development of approaches to need achievement, and the implications of this work for other aspects of motivation. This section is valuable to the student as a summary of one of the most active areas in the study of human motivation. The skeptical reader will feel that the coherence and agreement of the findings in this area have been overemphasized and that divergent findings have been underplayed.

Even taking the results on need achievement at face value as reported here, however, this reviewer finds that they have discouraging implications which do not seem to be recognized by Atkinson. Here we have a large body of research devoted to getting the beginnings of techniques for analyzing one motive. Yet even in the situations where this motive might be expected to be strong, the measures account for only a small part of the variance in performance and this is only one of a large number of motives still to be investigated and interrelated. The specification of this one motive seems to be unclear and variable even within the writings of the McClelland and Atkinson group of researchers. Does this mean that we must continue for many years in research designed to specify this one motive and devote comparable amounts of effort to each of the other motives which may be involved in behavior? McClelland and Atkinson would probably argue that the time devoted to need achievement is developmental, and that other motives can be more easily measured and studied as a result of this experience. This remains to be seen, however, and there is still another large task of finding the interrelations and interactions of all these motives once they have been (it is hoped) measured.

These discouraging prospects suggest to me that psychology must continue its search for other formulations and approaches to the problems of motivation, even if we accept the work of McClelland, Atkinson, and their colleagues as a valid approach. It may be valid but so inefficient that some overriding principles and methods are also urgently needed.

Returning to Cofer and Appley's contribution, we find its strength is an ability to summarize clearly a given theoretical position or a body of research upon a particular experimental problem. Its weakness lies in its over-all theoretical integration. Theories are presented side by side in insulated capsules; some very broad and abstract common features are sketched, but the theories are seldom confronted with each other. Criticisms of a particular theoretical position often consist largely of what other writers have said about the theory, and Cofer and Appley's own evaluation is not evident.

The authors own systematic views are given mainly at the beginning and end of the book, but the reviews also govern the organization and grouping of material in the rest of the book. The initial chapter introduces a conceptual vagueness which is also evident in later discussions. The title of Chapter I, "The concept of motivation" is a phrase which recurs, with the definite article, throughout the book. Yet the burden of this first chapter is that there are many different concepts which have been called motivational and which are used in accounting for behavior. Motivational terms "refer to concepts or constructs which are postulated to account for the behavior observed" (p. 9). Yet, in their final summing up, they say "We think it is evident from this discussion that motivation does not need to be invoked to account for the directional property of behavior or for the fact that behavior occurs. However, we do think that the motivation notion is needed to account for the invigoration of behavior" (p. 826). Nowhere between these two quotations can I find a definition of *the* motivational notion. Negatively, the motivational construct is not habit, innate structure, or ability, and it is not control by external stimuli. On the positive side, it seems to be some kind of "state" which provides stimuli—"If motivation does direct behavior, this can occur only through the stimuli which motivational states provide . . ." (p. 826). This does not seem to rule out very much, since it includes "symbolic and fantasied stimuli" (p. 828).

In contrast, Atkinson is very clear on what the field is all about, and this clarity gives greater coherence to his presentation. He begins with a common-sense statement of the *problem*, namely: "The study of motivation has to do with the contemporary (immediate) influences on the direction, vigor, and persistence of action" (p. 11). Some may want to change the scope of this definition (*e.g.* why limit it to action?), and most of the terms will need further precision as the study develops. Nevertheless, the field can be quite accurately defined without recourse to any particular set of explanatory constructs. Cofer and Appley's confusion is due to trying to define the problem in terms of some supposed characteristics of the solution to the problem. If this does not beg the question, it comes very close to it. Of course, constructs which are proposed as solutions to the problem fall into different classes, but it seems wasted effort to worry about which class is to be uniquely identified as "motivational." This attempt seems to go back to Hull's identification of motivation with his particular constructs of need and drive. Cofer and Appley reject drive as a useful construct, but seem to feel that they have to find something to take the place of drive as *the* motivational construct.

The vagueness of the notion of motivation does not detract from Cofer and Appley's book as much as might be expected. In effect, the book is about all of those concepts which have been traditionally grouped under the heading of motivation, and it can be read as a summary of a very large body of literature on these topics (some 1500 references) and within each topic the summaries are clear and informative.

In a final chapter, "Toward a unified theory of motivation", they concentrate upon three general ideas which they find useful in organizing the field. The first of these is the "equilibration model" or homeostasis. This certainly unifies some strange bedfellows, since it covers the points of view of G. L. Freeman, Helson, Stagner, Norbert Wiener, Ashby, and Lewin, as well as much of what has been subsumed under need or drive reduction.

The other two ideas are the *anticipation-arousal mechanism* (or *anticipation-invigoration mechanism*) and the *sensitization-invigoration mechanism.* The first of these covers the Hullian conception of secondary reinforcement, learned drives, and incentives treated as an alternative to drive. Formulations in terms of expectations are also considered analogous. In general, stimuli antedating or accompanying consummatory responses come to produce anticipation which in turn arouses or "invigorates" the organism. The same stimuli can also serve as cues to specific responses and may therefore be both arousing and directing. The sensitization-invigoration mechanism is designed to cover unlearned factors in arousal, such as the possibility that hormonal factors can have a selectively sensitizing effect for such things as sexual stimuli.

These concepts are considered distinct from the concept of drive, because Cofer and Appley consider drive as something which is dependent on the specific operation of deprivation. Drive in this sense they consider an unsatisfactory general concept, and point out inconsistencies and paradoxes which arise when drive is conceived in this manner, not only the difficulties with such things as exploratory or sexual behavior, but even with hunger which originally served as the model for the deprivation concept.

Thus Cofer and Appley have set themselves the task of teasing out what they can accept as a valid core of concepts from a vast array of different theories and types of research. What comes out seems to be so general as to have little predictive or explanatory power. Perhaps this is an inevitable result of trying to find common themes in a range of widely divergent theories and types of research all the way from the effects of food-deprivation in the white rat to the social behavior of the human adult. Such attempts are fostered by the bulk of theoretical writing and research formulations in the field, where the aim of finding all-inclusive principles dominates over the more modest goal of finding principles for a more limited range of phenomena. Yet there is a large gap between the all-inclusive principles and the specific interrelationships to be found for achievement behavior of the college student or the activity of the rat on short rations. If the study of motivation is ever to have any uses beyond furnishing topics for seminar discussions or themes of books for college students to read, the gap needs to be filled with principles which allow more specific predictions about behavior but are general enough to encompass more than one specific kind of response.

The tendency to seek extreme generality is also evident in Atkinson's treatment, in spite of his interest in need achievement, a motive with many practical implications for human behavior. His final summing up is devoted to showing the broad generality of the *Expectancy X Value* formulation, considering its application to the effects of food deprivation in rats, to avoidance conditioning in rats and dogs, and to the general formulations of the law of effect. In fact, he states explicitly ". . . the fundamental conception of motivation should apply to all instances of molar behavior" (p. 272), and "It is the ultimate or programmatic goal of all basic research on motivation to evolve a conceptual scheme that will provide a better understanding of human behavior" (p. 272), but he does not consider the possibility that there are intermediate schemes which might be more powerful in predicting specifically human behavior or certain classes or aspects of human behavior, leaving it to the future to determine how these principles could be subsumed under more general concepts applicable to other species.

Let me repeat that I am voicing here a complaint about a trend in the field of motivation at large which is reflected in these two books. Within this trend, both give us a wealth of information on the present status of the field.

Cornell University T. A. RYAN

Handbook of Mathematical Psychology. Edited by R. DUNCAN LUCE, ROBERT R. BUSH, and EUGENE GALANTER. New York, John Wiley and Sons, Inc. Vol. I & II—1963, Vol. III—1965. I, xiii, 491; II, vii, 606, III, ix, 537. $10.50, $11.95, $11.95.

Readings in Mathematical Psychology. Edited by R. DUNCAN LUCE, ROBERT R. BUSH, and EUGENE GALANTER. New York, John Wiley and Sons, Inc., 1963. Pp. v, 535. $8.95.

One is tempted to borrow adjectives from Hollywood with which to characterize these studies. They represent a colossal effort that brings us a stunning picture of the awesome confrontation in our century of mathematics and psychology. More soberly, the three-volume *Handbook* consists of 21 essays that survey mathematical thinking in 5 roughly-delineated areas of psychology: The psychophysics of perception, discrimination and choice, stochastic learning theory, mathematical methods, psycholinguistics, social psychology. The Volume I of *Readings* contains 35 articles reprinted from professional journals. Fourteen of these are classified by the editors as dealing with psychophysical topics, the remaining 21 as dealing with stochastic learning theory. Before undertaking a discussion of the scope and quality of these essays, it is important to consider the level of mathematical sophistication necessary to their understanding.

The editors tell us in the preface to Volume I of the *Handbook* that a reader without the equivalent of a solid bachelor's degree in mathematics may expect an uphill climb. Specifically, we are told (pp. vii, viii) that "the reader should . . . know foundational mathematics . . .; calculus; probability . . .; and statistics." Texts are suggested, such as Feller for probability, to indicate the level of depth needed. Even a cursory reading of the *Handbook* should convince one that these are altogether minimal requirements that must be considerably enriched to master some of the essays. Matrices and differential equations are ubiquitous. A course in the analysis of functions of real and complex variables is prerequisite to an understanding of at least a few of the articles. Stochastic process should be familiar in depth even greater than that of Feller's Volume I.

Psychophysics occupies seven chapters (2-5, 15, 16, 19) and about 35% of the *Handbook*. These include mathematical treatments of such traditional topics as auditory and visual perception, as well as newer ones that include signal detection, discrimination, and choice. In one of the *Handbook's* shorter essays, Chapter 2, the three editors collaborate to provide a classificatory scheme for choice-experiments that should be required reading for students of psychology. It is a small gem among taxonomic writings. Three chapters follow, all by Luce or by Luce and Galanter, on detection, recognition, discrimination, and psychophysical scaling. All lean heavily on Chapter 2 in their organization and notation. The four chapters comprise a unit that could be useful as an introductory psychophysics text. Chapter 15, "Analysis of some auditory characteristics," by J. Zwislacki, requires a solid grounding in auditory physiology and the mechanics of sound as well as in the fully panoply of mathematics. Chapter 15, by L. M. Hurvich, D. Jameson and D. H. Krantz, provides a less advanced, but nonetheless useful introduction to "Theoretical treatments of selected visual problems." It is unfortunate that neither chapter makes use of the editors' classificatory system. Chapter 19, by Luce and Suppes, comprises an exhaustive discussion (162 pages) of "Preference, utility, and subjective probability."

That there is considerable overlap between psychophysical and learning models is obvious and unavoidable. Still, five chapters of the *Handbook* (6, 9, 10, 17, 18), and about a quarter of its pages, seem to fit more in the area of learning than elsewhere. In Chapter 6, "Stochastic latency mechanisms," W. J. McGill discusses models of interresponse time-intervals in sequences of SRO experiments. Much of the excellence of this essay rests in its clarity of presentation and in its comparative analysis of competing models. In Chapter 9, Saul Sternberg writes at greater length about "Stochastic learning theory." Here, more and better data are brought to bear on the competing stochastic models than is the case in several of these essays. Atkinson and Estes, in one of the *Handbook's* longer chapters, give an outstandingly thorough account of stimulus sampling models. Once again, the primary tools are those of stochastic probability-processes. Comparisons of theory and data are frequent and encouraging. Algebraic treatments are straightforward, but unavoidably messy. Chapter 17, "Identification learning," by Robert R. Bush, follows closely on Chapters 9 and 10. The earlier of these chapters deals almost exclusively with learning behavior in two choice-situations. Bush generalizes the topic to include choice among *n* stimuli. His essay begins with a discussion of dual-process models, those of Bush and Mosteller, Restle, and Wyckoff in particular, that treat discrimination and conditioning as separate mechanisms. The remainder of the chapter is taken up with models of the single (conditioning) process in identification learning. This essay, rich in theorems and critical evaluation, is unsullied by data. Chapter 18, by Edwin Martin provides a brief overview of concept-utilization experiments, continuing Bush's consideration of the presentation of more than two stimuli. The summary analysis of concept-utilization as strategy-selection is especially valuable.

Three chapters (11-13) on linguistics comprise a 224-page work by Noam Chomsky (in collaboration with G. A. Miller on Chapters 11 and 13) that treats formal analyses of natural languages, grammars, and language-users. None of the chapters is especially demanding mathematically, but some familiarity with linguistics, graphs and information-theory is prerequisite to understanding them.

Five chapters (1, 7, 8, 20, 21) that are more purely methodological than the others are uniformly excellent, and some are superb. Suppes and Zinnes provide, in Chapter 1, a definitive treatment of measurement-theory. It is a demanding essay that requires some sophistication in formal logic, but if its demands can be met, its rewards are considerable. In Chapter 7, Allen Newell and Herbert Simon discuss a number of computer-applications in psychology. Their description is lucid and useful, yet nearly devoid of mathematical content. As the authors point out, the inclusion of such a chapter in the *Handbook* is nothing more than an historical accident. They discuss simulation in psychology in some depth, but, regrettably, their essay has nothing to say about on-line computer-uses in learning and interaction-research. Robert Bush follows with a short paper on estimation- and evaluation-procedures in learning and psychological experiments. Chapter 20 consists of an introduction to stochastic processes by J. L. Snell. Topically, it is intermediate to the Dartmouth Group's *Finite Markov Chains* and its recently-published *Denumerable Markov Chains*. Snell discusses both classes of chains, each in both discrete and continuous time. The chapter provides a more comprehensive introduction to stochastic processes than does any of the similar works of the Dartmouth Group.

The single chapter that is given over to social psychology (Chapter 14, by A. Rapaport) is disappointing both in its brevity and in its selectivity. It is not surprising to find detailed treatments of net structures, dominance-theory, and prisoner's dilemma games, all of which are closely associated with the author's name. It is unfortunate, however, to find such topics as coalition-formation, conformity-models, and group problem-solving either relegated to short paragraphs or neglected altogether.

A number of potshots could be taken at the contents of the *Handbook*. The quality and scope of its essays vary markedly, as do the demands that are placed on its readers. The writing is rarely exciting and occasionally tedious (Bush, Luce, and McGill provide engaging exceptions). The almost random location of articles that should be grouped topically is irritating. To fire off any more than the shortest volley, however, would be ridiculous. The three volumes of the *Handbook* stand and will not be shot down. Surely they will be recognized as a milestone of twentieth century psychology.

Cornell University

ROBERT MCGINNIS

Social Psychology. By ROGER BROWN. New York, Free Press, 1965. Pp. xxiv, 785. $8.95.

Roger Brown's *Social Psychology* differs from most books on that subject in that it is neither a textbook that presents one systematic point of view (psychological or sociological), nor a compendium of facts distilled from reports of research. It is written, not by a psychologist and a sociologist collaborating to present the best of their respective disciplines amicably, but by one man who takes full responsibility for what he says. What he calls "fourteen largely independent chapters" are essentially essays on "research traditions" in social psychology, all scholarly and thoughtful, varying in depth and threat of difficulty for the undergraduate, and in style from chatty remarks about life at Harvard and MIT to relatively profound discussions of theory and research. He disclaims any effort to integrate the topics or even to employ a common language in discussing them.

"Social psychology in 1964 is a set of topics that have exceeded the grasp of a non-social psychology but which are being effectively investigated by a psychology that draws upon the social sciences" (pp. xix-xx).

"Professor Brown begins by familiarizing the student with the basic concepts and principles of social psychology," says the dust jacket. Neither the Preface nor the Introduction did that to me. Nor did the first chapter, The Social Behavior of Animals, unless animal population-adjustment, starvation as a regulating mechanism, territoriality, imprinting, and the monkey infant's preference for a terrycloth mother-surrogate constitute these "basic concepts and principles." As proof that "the study of animal behavior also exerts a force on our conceptions of ourselves and our species," Brown jocularly cites Kay, in Mary McCarthy's novel *The Group,* whose life had been amazingly altered by a course in Animal Behavior she had taken with old Miss Washburn (who had left her brain in her will to Science)" (p. 42).

In Part II, Problems of Psychology and Social Structure, Brown effectively employs an interesting account of the evolution of European pronouns of address to introduce the more conventional concepts of status, norms, social stratification, roles, and stereotypes.

In Part III, The Socialization of the Child, the six chapters are more interrelated and inter-relevant than the author led us to believe they would be. Introducing the language system and its acquisition by way of Piaget's work on the growth of intelligence in the child, Brown presents a relatively detailed account of the development of phonology and grammar, including data from observations of children by a research group to which he belonged. In discussing the semantic system, he draws on several well-known studies of concept development. He explains the codability of referents mainly in terms of Piaget's model of the child's egocentrism in speech and thinking. In his account of the acquisition of morality, he ranges from Freud's *Analysis of a Phobia in a Five-Year-Old Boy* and Piaget's studies of moral judgment, through Hartshorne and May, Sears, Maccoby and Levin, Whiting, MacKinnon, Bandura *et al,* to the more recent contributions of Kohlberg at Chicago. Here, I think, Brown has gone a long way toward establishing *learning* as the proper basis for understanding and predicting the development and acquisition of culture. Abandonment of the concept of *interiorization* of norms (or culture) in favor of *learning* as explanatory principle, it seems to me, is long verdue, and Brown's contributions represent real progress toward this goal.

Part IV, Personality and Society, concerns mainly the Achievement Motive and the Authoritarian Personality. Reviewing the controversies over interpretation of the latter, Brown concludes that "it has not been demonstrated that fascists and communists resemble one another in authoritarianism or in any other dimensions of ideology" (p. 542).

Part V, Social Psychological Processes, returns to more conventional discussion of consistency in attitude change, impressions of personality, group dynamics, collective behavior, and the psychology of the crowd. The chapter on the latter two topics, which includes the Iroquois Theater fire, "tulipomania" in Holland, and lynchings that the author has discussed elsewhere, is the most striking of them all. If such reports do no more than sharpen our observation of crowd phenomena in these parlous times, they serve a good purpose. Here, as elsewhere, Brown's language occasionally takes a somewhat archaic "mentalistic" turn: "a social unit will not do the job. The locus involved (perhaps it is a substance) must really be the human mind" (p. 732), "conditions apprehended by the same mind" (p. 732), "social controls and the rest are reinterpreted as mental forces" (p. 733). Considering Brown's academic background (Michigan, Marvard, MIT), it seems strange that the chapter on Group Dynamics contains little or no reference to Lewin, J. R. P. French, Bavelos, or Ronald Lippitt, being based largely on the balance model and decision-making. Little attention is paid to perception (other than perception of persons), to persuasive communication, social power, leadership, or group cohesiveness. With hardly an attempt to define the general nature of attitudes, discussion of that topic is limited mainly to attitude change.

In substance, the jacket's claim that "this book treats all the major topics and notable experiments in social psychology with an excellent balance between early and recent research," falls short of fulfillment. Within the limits Brown set for himself, however, about the only faults I can find are a few errors in names of authors that might trouble students bent on looking up references. For a busy instructor seeking what the jacket promises, Brown's self-imposed limits on subject matter might give him pause. Brown even suggests omitting some chapters because he himself "has never been able to include all these topics in a single semester course" (p. vii). On the credit side, I believe the fourteen chapters are more integrated and unified than Brown knows. What holds the book together, it seems to me, is the style of the man and the frankness with which he shares his thinking.

North Carolina Wesleyan College ARTHUR JENNESS

Personnel Testing. By ROBERT M. GUION. New York, McGraw Hill, 1965. Pp. xiii, 585. $9.50.

Guion has done a remarkably good job of integrating the relevant literature concerning personnel selection, evaluating it, and pointing the direction in which more innovative research should go. This book has been needed for a long time. It was a pleasure for this reviewer to use it as a text last term in an undergraduate course. The students found it difficult but challenging, and commented that it was so well organized and written that they found the task of estimating the relative importance of the materials presented an unusually easy one.

This book is required reading for psychologists in industrial, psychometric, or other measurement areas, and should even clarify the thinking of many in other areas as well.

Cornell University PATRICIA CAIN SMITH

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

ABRAHAMSON, MARK, *Interpersonal Accommodation.* Princeton, N.J., D. VanNostrand, 1966. Pp. vii, 151. $1.75 (paperback).

AMOS, WILLIAM E., and MANELLA, RAYMOND L., (eds.), *Delinquent Children in Juvenile Correctional Institutions. State Administered Reception and Diagnostic Centers.* Springfield, Ill., C. C. Thomas, 1966. Pp. xiii, 159. $7.50.

BAKAN, PAUL (ed.), *Attention. An Enduring Problem in Psychology.* Princeton, N. J., D. VanNostrand, 1966. PP. vi, 225. $1.75 (paperback).

BAUGHMAN, LAURENCE ALAN, *Southern Rape Complex. Hundred Year Psychosis.* Atlanta, Ga., Pendulum Books, 1966. Pp. 222. $2.95 (paperback).

BERGEN, BERNARD J., and THOMAS, CLAUDEWELL S., *Issues and Problems in Social Psychiatry. A Book of Readings.* Springfield, Ill., C. C. Thomas, 1966. Pp. xi, 468. $14.50.

BETTELHEIM, BRUNO, *The Empty Fortress. Infantile Autism and the Birth of the Self.* New York, Free Press, 1967. Pp. xiv, 484. $9.95.

BINDRA, DALBIR, and STEWART, JANE (eds.), *Motivation. Selected Readings.* Baltimore, Md., Penguin Books, 1966. Pp. 352. $1.95 (paperback).

BIRAN, SIGMUND, *Die ausserpsychologischen Voraussetzungen der Tiefenpsychologie.* München-Basel, Ernst Reinhardt Verlag, 1966. Pp. 180. DM 11.

BOLLES, ROBERT C., *Theory of Motivation.* New York, Harper & Row, 1967. Pp. vii, 546. $10.75.

BRESSON, F., and MONTMOLLIN, M. (eds.), *Psychologie et Epistemologie Genetiques. Themes Piagetiens. Hommage a J. Piaget avec une bibliographie complete de ses oeuvres.* Paris, Dunod, 1966. Pp. xxi, 421. 39 F (paperback).

CHAMBLISS, ROLLIN, *Meaning for Man.* New York, Philosophical Library, 1966. Pp. xi, 191. $4.50.

CHAPMAN, J. DUDLEY, *The Feminine Mind and Body.* New York, Philosophical Library, 1967. Pp. 325. $6.95.

COLE, JONATHAN O., and WITTENBORN, J. R. (eds.), *Pharmacotherapy of Depression.* Springfield, Ill., C. C. Thomas, 1966. Pp. xi, 189. $8.50.

CRAFT, M. (ed.), *Psychopathic Disorders and Their Assessment.* Long Island City, N.Y., Pergamon Press, 1966. Pp. x, 320. $8.00.

DEESE, JAMES, and HULSE, STEWART H., *The Psychology of Learning.* 3rd ed. New York, McGraw-Hill, 1967, Pp. viii, 514. $7.95.

DE JOUVENEL, BERTRAND, *The Art of Conjecture.* (Translated by Nikita Lary) New York, Basic Books, 1967. Pp. xii, 307. $7.50.

DE LAGUNA, GRACE A., *On Existence and the Human World.* New Haven, Conn., Yale Univ. Press, 1966. Pp. x, 267. $6.50.

EPSTEIN, WILLIAM, *Varieties of Perceptual Learning.* New York, McGraw-Hill, 1967. Pp. vii, 323. $9.50.

ERON, LEONARD D. (ed.), *The Classification of Behavior Disorders.* Chicago, Ill., Aldine Pub. Co., 1967. Pp. xii, 180. $5.95.

EXNER, JOHN E., *A Workbook in the Rorschach Technique Emphasizing the Beck and Klopfer Systems.* Springfield, Ill., C. C. Thomas, 1967. Pp. viii, 113. $10.50.

FARBER, SEYMOUR M., WILSON, NANCY L., and WILSON, ROGER H. L. (eds.), *Food and Civilization. A Symposium.* Springfield, Ill., C. C. Thomas, 1966. Pp. xii, 341. $11.75.

FELDMAN, SHEL (ed.), *Cognitive Consistency. Motivational Antecedents and Be*havioral Consequents. New York, Academic Press, 1966. Pp. xiii, 312. $8.50.

FRANKEL, MAX G., HAPP, F. WILLIAM, and SMITH, MAURICE P., *Functional Teaching of the Mentally Retarded.* Springfield, Ill., C. C. Thomas, 1966. Pp. xvii, 241. $9.75.

Gellhorn, Ernst, *Principles of Autonomic-Somatic Integrations. Physiological Basis and Psychological and Clinical Implications.* Minneapolis, Minn., Univ. of Minn. Press, 1967. Pp. xiii, 318. $10.00.

Gordon, Jesse E. (ed.), *Handbook of Clinical and Experimental Hypnosis.* New York, Macmillan, 1967. Pp. viii, 653. $14.95.

Group for the Advancement of Psychiatry, *Psychiatric Research and the Assessment of Change.* Report No. 63. New York, Group for the Advancement of Psychiatry, 1966. Pp. 351-478. $2.50 (paperback).

Guilford, J. P. (ed.), *Fields of Psychology.* 3rd ed. Princeton, N.J., D. VanNostrand, 1966. Pp. x, 350. $8.95.

Halloran, J. D. *Attitude Formation and Change.* Leicester, England, Leicester Univ. Press, 1967. Pp. 167. 21s (paperback).

Hellmuth, Jerome (ed.), *Educational Therapy. Vol. 1.* Seattle, Wash., Special Child Publications of the Seattle Seguin School, Inc., 1966. Pp. 468. $10.00.

Himwich, Harold E., Kety, Seymour S., and Smythies, John R. (eds.), *Amines and Schizophrenia.* Long Island City, N.Y., Pergamon Press, 1967. Pp. ix, 290. $11.50.

Hoffman, Lois Wladis, and Hoffman, Martin L. (eds.), *Review of Child Development Research. Prepared Under the Auspices of the Society for Research in Child Development.* Vol. 2. New York, Russell Sage Foundation, 1966. Pp. xi, 598. $8.00.

Hope, K., *Elementary Statistics. A Workbook.* Long Island City, N.Y., Pergamon Press, 1967. Pp. xii, 101. $4.00 (paperback).

Hurvich, Leo M., and Jameson, Dorothea, *The Perception of Brightness and Darkness.* Boston, Allyn and Bacon, 1966. Pp. ix, 141. $2.50 (paperback).

Inglis, James, *The Scientific Study of Abnormal Behavior.* Chicago, Ill., Aldine Pub. Co., 1967. Pp. xiv, 256. $8.95.

Jahoda, Marie, and Warren, Neil (eds.), *Attitudes. Selected Readings.* Baltimore, Md., Penguin Books, 1966. Pp. 375. $1.95 (paperback).

Jeffery, C. R., *Criminal Responsibility and Mental Disease.* Springfield, Ill., C. C. Thomas, 1967. Pp. xvii, 324. $11.00.

Jourard, Sidney M., and Overlade, Dan C. (eds.), *Reconciliation. A Theory of Man Transcending. From the Work of Franklin J. Shaw.* Princeton, N.J., D. VanNostrand, 1966. Pp. x, 148. $1.75 (paperback).

Jung, C. G., *Collected Works, Vol. 7. Two Essays of Analytical Psychology.* Edited by Gerhard Adler, Michael Fordham, and Herbert Read. 2nd ed. New York, Pantheon Books, 1966. Pp. xi, 357. $4.50.

Jung, C. G., *Collected Works, Vol. 15. The Spirit in Man, Art, and Literature.* Edited by Gerhard Adler, Michael Fordham and Herbert Read. New York, Pantheon Books, 1967. Pp. viii, 168. $4.50.

Kaelbling, Rudolph, and Patterson, Ralph M., *Eclectic Psychiatry.* Springfield, Ill., C. C. Thomas, 1966. Pp. xxv, 891. $17.50.

Katz, Jay, Goldstein, Joseph, and Dershowitz, Alan M., *Psychoanalysis, Psychiatry and Law.* New York, Free Press, 1967. Pp. xxiii, 822. $15.00.

Keve, Paul W., *Imaginative Programming in Probation and Parole.* Minneapolis, Minn., Univ. of Minnesota Press, 1967. Pp. 293. $6.75.

Klausmeier, Herbert J., and Harris, Chester W. (eds.), *Analyses of Concept Learning.* New York, Academic Press, 1966. Pp. xiii, 272. $8.50.

Kreitler, Shulamith, *Symbolschöpfung und Symbolerfassung. Eine experimentalpsychologische Untersuchung.* München-Basel, Ernst Reinhardt Verlag, 1965. DM 20.

Lawson, Chester A., *Brain Mechanisms and Human Learning.* Boston, Mass., Houghton Mifflin, 1967, Pp. xiii, 126. $4.50.

Lenneberg, Eric H., *Biological Foundations of Language.* New York, Wiley, 1967. Pp. xvi, 489. $14.95.

Lewis, Weldon, *Air and Its Function in Life.* New York, Vantage Press, 1966. Pp. 64. $2.50.

Lindvall, C. M., *Measuring Pupil Achievement and Aptitude.* New York, Harcourt, Brace & World, 1966. Pp. xi, 188. $2.25 (paperback).

LIPMAN, MATTHEW, *What Happens in Art.* New York, Appleton-Century-Crofts, 1967. Pp. vii, 177. $1.95 (paperback).

LOUCH, A. R., *Explanation and Human Action.* Berkeley, Calif., Univ. of California Press, 1966. Pp. viii, 243. $6.95.

LYONS, J., and WALES, R. J. (eds.), *Psycholinguistics Papers. Proceedings of the Edinburgh Conference 1966.* Chicago, Ill., Aldine Pub. Co., 1967. Pp. 243. $7.95.

MANDLER, GEORGE; MUSSEN, PAUL; KOGAN, NATHAN and WALLACH, MICHAEL A., *New Directions in Psychology III.* New York, Holt, Rinehart & Winston, 1967. Pp. ix, 289. $3.95 (paperback).

MARTIN, GUSTAV J., and KISCH, BRUNO (eds.), *Enzymes in Mental Health.* Philadelphia, Pa., Lippincott, 1966. Pp. 208. $8.00.

MAY, ROLLO, *Psychology and the Human Dilemma.* Princeton, N.J., D. VanNostrand, 1966. Pp. iii, 221. Cloth, $5.95; paper, $2.95.

MCCLELLAND, DAVID, *The Achieving Society.* New York, Free Press, 1967. (First published, 1961). Pp. xv, 512. $2.95 (paperback).

MCGAUGH, JAMES L., WEINBERGER, NORMAN M., and WHALEN, RICHARD E., *Psychobiology. The Biological Bases of Behavior. Readings from Scientific American.* San Francisco, Calif., W. H. Freeman, 1967. Pp. 382. Cloth, $10.00; paper, $5.45.

MCKINNEY, FRED (ed.), *Psychology in Action. Basic Readings.* New York, Macmillan, 1967. Pp. xii, 564. $3.95.

MELTZOFF, JULIAN, and BLUMENTHAL, RICHARD L., *The Day Treatment Center. Principles, Application and Evaluation.* Springfield, Ill., C. C. Thomas, 1966. Pp. xvi, 234. $8.75.

MOSS, C. SCOTT, *The Hypnotic Investigation of Dreams.* New York, Wiley, 1967. Pp. xi, 290. $7.50.

O'CONNOR, NEIL (ed.), *Present-day Russian Psychology. A Symposium by Seven Authors.* Long Island City, N.Y., Pergamon Press, 1966. $6.50 (paperback).

PACIFICO, CARL, Creative Thinking in Practice. Park Ridge, N.J., Noyes Press, 1966. Pp. 114. $9.00.

PELZ, DONALD C., and ANDREWS, FRANK M., *Scientists in Organizations. Productive Climates for Research and Development.* New York, Wiley, 1966. Pp. xii, 318.

PIAGET, JEAN, *Six Psychological Studies.* (Translated by A. Tenzer and D. Elkind) New York, Random House, 1967. Pp. xvii, 169. $4.95.

RAY WILBERT S., *The Experimental Psychology of Original Thinking.* New York and London, Macmillan, 1967. Pp. xiii, 208. $1.75 (paperback).

RAYCHAUDHURI, MANAS, *Studies in Artistic Creativity. Personality Structure of the Musician.* Calcutta, India, Rabindra Bharati, 1966. Pp. xii, 256.

RHINE, LOUISA E., *ESP in Life and Lab. Tracing Hidden Channels.* New York, Macmillan, 1967, Pp. viii, 275. $6.95.

ROTHWELL, NAOMI D., and DONIGER, JOAN M., *The Psychiatric Halfway House. A Case Study.* Springfield, Ill., C. C. Thomas, 1966. PP. xvi, 265. $8.75.

SAWYER, PHILIP L., *A Neural Systemic Theory of Emotion. An Outline of a New Methodological Approach to Psychology and a Theory of Emotion and the Mind.* Brooklyn, N.Y., Theo Gaus' Sons, 1966. Pp. xii, 57. Cloth, $2.95; paper, $1.75.

SCIENTIFIC AMERICAN (eds.), *A Comprehensive Review of the Extraordinary New Technology of Information.* San Francisco, Calif., W. H. Freeman, 1966. Cloth, $5.00; paper, $2.50.

SEGALL, MARSHAL H., CAMPBELL, DONALD T., and HERSKOVITS, MELVILLE J., *The Influence of Culture on Visual Perception. An Advanced Study in Psychology and Anthropology.* Indianapolis, Ind., New York, and Kansas City, Mo., Bobbs-Merrill, 1966. Pp. xvii, 268. $2.95 (paperback).

SEMEONOFF, BORIS (ed.), *Personality Assessment. Selected Readings.* Baltimore, Md., Penguin Books, 1966. Pp. 443. $1.95 (paperback).

SHAW, MARVIN E., and WRIGHT, JACK M., *Scales for the Measurement of Attitudes.* New York, McGraw-Hill, 1967. Pp. xxii, 604. $14.50.

SMART, MOLLIE S., and SMART, RUSSELL C., *Child Development and Relationships.* New York, Macmillan, 1967. Pp. vii, 582. $7.95.

SPROTT, W. J. H., *Social Psychology.* New York, Barnes & Noble, 1967. (First published, 1952). Pp. xiv, 268. Cloth, $6.25; paper, $3.50.

STEKEL, WILHELM, *Frigidity in Woman in Relation to Her Love Life.* New York, Washington Square Press, 1967. Pp. xi, 627. $1.45 (paperback).

STEPHENSON, WILLIAM, *The Play Theory of Mass Communication.* Chicago, Ill., Univ. of Chicago Press, 1967. Pp. viii, 225. $5.00.

STONE, PHILIP J., DUNPHY, DEXTER C., SMITH, MARSHALL S., OGILVIE, DANIEL M., ET AL, *The General Inquirer: A Computer Approach to Content Analysis.* Cambridge, Mass., and London, Engl., M.I.T. Press, 1966. Pp. xx, 651. $7.95.

TELFORD, CHARLES W., and SAWREY, JAMES M., *The Exceptional Individual. Psychological and Educational Aspects.* Englewood Cliffs, N.J., Prentice-Hall, 1967. Pp. x, 482. $7.95.

TRAVERS, ROBERT M. W., *Essentials of Learning. An Overview for Students of Education.* 2nd ed. New York, Macmillan, 1967. Pp. xi, 560. $7.95.

VAN DEN BERG, J. H. *The Psychology of the Sickbed.* Pittsburgh, Pa. Duquesne Univ. Press, 1967. Pp. 136. $2.95.

VERNON, M. D. (ed.), *Experiments in Visual Perception. Selected Readings.* Baltimore, Md., Penguin Books, 1966. Pp. 447. $1.95 (paperback).

VON FIEANDT, KAI, *The World of Perception.* Homewood, Ill., Dorsey Press, 1966. Pp. ix, 418.

WAHL, CHARLES WILLIAM (ed.), *Sexual Problems. Diagnosis and Treatment in Medical Practice.* New York, Free Press, 1967. Pp. xix, 300. $7.95.

WHITE, MARY ALICE, and CHARRY, JUNE (eds.), *School Disorder, Intelligence and Social Class.* New York, Teachers College Press, Columbia Univ., 1966. Pp. viii, 92. $2.25 (paperback).

WITTENBORN, J. R., *The Clinical Psychopharmacology of Anxiety.* Springfield, Ill., C. C. Thomas, 1966. Pp. x, 228.

ZIMMER, HERBERT (ed.), *Computers in Psychophysiology.* Springfield, Ill., C. C Thomas, 1966. Pp. xiii, 125. $8.50.

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

VOL. LXXX SEPTEMBER, 1967 NO. 3

ADAPTATION-LEVEL THEORY: TWO CRITICAL EXPERIMENTS ON HELSON'S WEIGHTED-AVERAGE MODEL

By VIKTOR SARRIS, University of Duesseldorf, Germany

During the last twenty years, Helson, together with his co-workers, has propounded and developed the theory of "adaptation-level" (*AL*).[1] This comprehensive psychological theory of relativity is based on the general principle of perceptual contrast, and it has been formulated in precise mathematical terms. The most essential part of the definition is the weighted log mean which has been used till now mainly for psychophysical data but is apparently not confined to them.[2] Its relevant *mathematical* implications, however, have never been sufficiently tested and developed by systematic means, *i.e.* trend-analytical experimentation.

Helson's basic equation for *AL* has the following form:

$$A = K(\overline{S}^p C^q R^r), \tag*{[1]}$$

where A is the predicted value of AL, K an empirical constant, $\overline{S}$ the geometric mean of the series-stimuli, C the anchor-stimulus, R refers to residual stimuli, and

---

* Received for publication January 20, 1967. Based on a thesis submitted in partial fulfillment of the requirements for the Doctoral degree at the University of Hamburg, Germany. The results of this study have been described and discussed in greater detail elsewhere (see Viktor Sarris, Die Abhängigkeit des Adaptationsniveaus von Ankerreizen, *Z. exp. angew. Psychol.*, 14, 1967, 1-53. Special thanks are due to Dr. G. A. Lienert, University of Duesseldorf, for supervision of the experiments.

[1] Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29; Adaptation-level as a base for a quantitative theory of frames of reference, *Psychol. Rev.*, 55, 1948, 297-313; Adaptation level theory, in Sigmond Koch (ed.), *Psychology: A Study of a Science: I, Sensory, Perceptual, and Physiological Foundations*, 1959, 565-621; *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior*, 1964.

[2] See J. E. Podell, A comparison of generalization and adaptation-level as theories of connotations, *J. abnorm, soc. Psychol.*, 62, 1961, 593-597; cited also in Helson, *op. cit.* 1964, 560 ff.

p, *q*, and *r* are constant weighting coefficients which must be determined empirically and the numerical sum of which is taken to be 1.0.[3] This is, however, only the *general* formula for *AL*. Beyond it, taking into account the most frequent experimental case, Helson has gone on to assume that the residual stimuli are practically irrelevant and has also replaced the constant K by an additive constant $c \cdot d$ (the so-called "*d*-factor"). Doing this, Helson got a formula whose logarithmic form now reads:

$$\log(A + c \cdot d) = \frac{k \log \bar{S} + \log C}{k + 1}, \qquad [2]$$

where k is the empirically determined constant weighting coefficient for the series-stimuli (S) in relation to an anchor-stimulus (C), d is the step-interval between the series-stimuli, and c is the empirical value of time-order effects produced as the result of *successive* presentation of series-stimuli and the anchor-stimulus.[4] If in a particular experiment *no* anchor-stimulus is employed, the term C in Equation [2] drops out, and Helson predicts the *AL* as follows:

$$\log\,(A + c \cdot d) = \log \bar{S} \qquad [3]$$

Empirically, *AL* is defined as the stimulus which elicits a neutral response. By using Helson's methods for estimating the *empirical AL*-values,[5] it becomes possible to test the empirical validity of the Equations [2] and [3] for different psychophysical or other materials.

Though Helson and his co-workers have tried to demonstrate the usefulnesss of the quantitative *AL*-model as to various psychological phenomena by means of many experiments,[6] and though several eminent psychologists have appreciated the *AL*-model both because of its quantitative foundation and its theoretical versatility,[7] *AL*-theory has nevertheless also been criticized by some writers for different reasons.[8]

---

[3] Cf. Helson, *op. cit.*, 1964, 58 and 129. In presenting Helson's basic equations here, substitutions for some of the symbols have been made in order to facilitate exposition in the present context.

[4] Cf. Helson, *op. cit.*, 1964, 129-132.

[5] Harry Helson and Philip Himelstein, A short method for calculating the adaptation-level for absolute and comparative rating judgments, this Journal, 68, 1955, 631-637; see also Helson, *op. cit.*, 1948, 302-307; *op. cit.*, 1964, 661-662.

[6] See Helson, *op. cit.*, 1964.

[7] See *e.g.* Floyd Allport, *Theories of Perception and the Concept of Structure*, 1955, 242 ff.; J. P. Guilford, *Psychometric Methods*, (2nd ed.), 1954, 327 ff.; With scope and relevance, *Contemp. Psychol.*, 10, 1965, 241-242, (Book review of Helson, *op. cit.*, 1964); P. R. Hofstätter, *Sozialpsychologie*, (2nd ed.) 1966, 130 ff.; Wilhelm Witte, Das Problem der Bezugssysteme, in Kurt Gottschaldt, Philipp Lersch, and Hans Thomae (eds.). *Handbuch der Psychologie, (12 Bände), I, 1, Erkennen und Wahrnehmen*, 1964, 1013-1021.

[8] See *e.g.* D. R. Brown, Stimulus-similarity and the anchoring of subjective scales, this Journal, 66, 1953, 199-214; Gloria Engel and Allen Parducci, Value of background in the specification of the stimulus for judgment, this Journal, 74, 1961, 569-575; H. W. Hake, Perception, *Annu. Rev. Psychol.*, 13, 1962, 145-170; David Krech and R. S. Crutchfield, *Elements of Psychology*, 1958, 95; Allen Parducci, Range-frequency compromise in judgment, *Psychol. Monogr*, 77, 1963, (No. 565), 1-50; Adaptation-level theory, this Journal, 78, 1965, 158-159 (book review of Helson, *op. cit.* 1964); S. S. Stevens, Adaptation-level vs. the relativity of judgment, this Journal, 71, 1958, 633-646; J. F. Wohlwill, Perceptual learning, *Annu. Rev.*, 17, 1966, 201-232.

Among the skeptical views against Helson's theory, Parducci's argument is the most essential one, namely:

> "The psychophysical and perceptual applications should be approached as *illustrations* of the theory rather than as quantitative tests. Although the range of experimental situations is impressive, the data from any specific situation are too sparse to provide more than a directional or qualitative test. . . . Considered as a quantitative theory, its evaluation must await further development and testing."[9]

The results of a preliminary *AL*-study concerned with the systematic effects of different anchor-stimuli upon the judgment of size (circles) agree with this critical statement.[10] For it was shown in this investigation that it was impossible to predict the established anchor-*AL* relationship by use of Helson's Equation [2]. Far from being a "constant," the weighting coefficient k varied with the size of anchor-stimuli.

This result, together with the following remark by Helson himself, raised some concrete doubt on the empirical validity of the quantitative *AL*-formula. For Helson—in contrast to his mathematical implication that with constant series and variable anchor-stimulus *AL* varies as the $(k + 1)$. root of anchor size[11]—has also stated:[12] "We ourselves do not believe that *AL* is monotonically related to the values of the anchors except within certain limits."[13]

These "certain limits" are, in principle, not predictable from Helson's basic formulae. Now, since the *quantitative* essence of whole *AL*-theory is based on the mathematical structure of Helson's fundamental formula, a *crucial experiment* concerning the predictive value of Helson's formula or formulae in question (see above) seemed to be indispensable. For if one is really to presuppose the *general* validity of the basic quantitative *AL*-model, a proof of its accuracy as to Helson's own original experimental procedure may, at least, be demanded. Therefore, in order to test the fundamental *AL*-formula empirically as adequately as possible, it was decided for this critical investigation to use a strict *replication* of Helson's experimental arrangement used in 1947 for lifted-weights.[14]

In Equations [2] and [3], there are involved two different mathematical relationships the experimental investigation of which seems to be most relevant for the quantitative theory, namely:

(1) The anchor-*AL*-relationship in Equation [2]; for Helson's so-called "pooling" process by the organism in respect to variables stemming from different stimuli classes plays an important role in *AL*-theory.[15]

---

[9] Parducci, *op. cit.*, 1965, 158 f. (Italics ours.)

[10] Unpublished research work by the writer carried out in autumn 1964 and cited in Viktor Sarris, *op. cit.*, 14, 1967, 11 f.

[11] Helson, *op. cit.*, 1947, 5; also Cf. Helson, *op. cit.* 1964, 131.

[12] Personal communication. Professor Helson's letter to the author dated September 15, 1964 is kindly acknowledged.

[13] Actually, this statement has been experimentally verified by Helson and Masters, but interestingly enough the authors did not revise the basic mathematical *AL*-model according to their empirical results (see Harry Helson and H. G. Masters, A study of inflection-points in the locus of adaptation-levels as a function of anchor-stimuli, this JOURNAL, 79, 1966, 400-408; see also Sarris, Comments on Helson and Masters' paper "A study of inflection-points in the locus of adaptation-levels as a function of anchor-stimuli", this JOURNAL, 80, 1967, 304-309.

[14] Helson, *op. cit.*, 1947, 7 ff.

[15] Helson, *op. cit.*, 1964, 73-88.

(2) The step-interval-*AL*-relationship in Equation [3], also implied in Equation [2]; for Helson has never exactly proved the theoretical and empirical justification of his '*d*-factor.' This point has been also stated by Parducci: "For some data he (Helson) adds an arbitrary correction-factor (0.75*d*) which seems out of character with his theory of stimulus-pooling."[16]

In 1947, with a constant stimulus-series consisting of five weights (200, 250, 300, 350, and 400 gm.), Helson has experimentally established the values of the empirical constants in his basic Equations [2] and [3], namely: $k = 3.0$ and $c = 0.75$.[17] Accordingly, in the present study the empirical validity of the following two Equations was investigated:

(1) Experiment I *with* anchor-stimulus involved:

$$A = 70.5\ C^{1/4} - 37.5 \qquad [2a]$$

This equation originates from use of Helson's above values for Equation [2].

(2) Experiment II with *no* anchor-stimulus involved:

$$A = 291.4 - 0.75\ d \qquad [3a]$$

This relationship is equivalent to Equation [3], provided Helson's experimental values are used, respectively.

EXPERIMENT I: ANCHOR-*AL* RELATIONSHIP

Method: (1) Stimuli. To replicate Helson's experimental procedure of 1947, a stimulus-series consisting of 200, 250, 300, 350, and 400 gm. was produced. For this, five plastic bottles of 500 ccm. volume each, measuring 13.0 cm. in height and having a circular bottom with 6.5 cm. in diameter, were precisely filled with respective amounts of lead shot.[18] In addition to these stimulus-objects, there were ten other bottles of exactly the same size but of different weights which served as the anchor stimuli. These anchor weights had the following values: 12.4, 35, 75, 291.4, 450, 900, 1500, 2400, 3400, and 4500 gm. These values forming an irregular series were chosen deliberately in order to test Helson's *"certain limits"* in question (see above) both as safely and economically as possible in respect to the practically relevant anchor-*AL* relationship and to the experimental expense. The anchor weight of 291.4 gm. served as an experimental control condition in that according to Helson's formula [2] this anchor-stimulus should produce *no* contrast effect.[19]

(2) Apparatus. To present the stimuli to the *S*s as precisely and quickly as feasible, a semi-automatic apparatus was used. This apparatus functioned like a 'telephone dial' and was attended by *E*, who had to present the single weights successively to *S*'s hand, always into the same lifting position. The apparatus (its

---

[16] Parducci, *op. cit.*, 1965, 158 f.; see also Sarris, *op. cit.*, *Z. exp. angew. Psychol.*, 14, 1967, 10 f.

[17] Helson, *op. cit.*, 1947, 4 f.; *op. cit.*, 1964, 130 f.

[18] Professor Helson's letter to the present writer containing particular explanations as to the original experimental procedure from 1947 is kindly acknowledged (personal communication, August, 1965).

[19] A critical remark concerning Helson's own "control" anchor-weight (cf. Helson, *op cit.*, 1947, 9; also 1964, 133) is given in Sarris, *op. cit.*, *Z. exp. angew. Psychol.* 1967, 14, 15.

main part being a turn-table) and stimuli were concealed from *S* by a wooden screen which had only a small hole at its right lower side for *S*'s right hand. So the weights were never visible to *S*.[20]

(3) Subjects. The *S*s were right-handed soldiers, their ages ranging between 19 and 24 yr. The *S*s (160), who were not familiar with the problem being investigated, were randomly assigned to the 10 different anchor conditions (see above) so that each experimental group consisted of 16 *S*s and had to judge under only *one* experimental condition ("randomized groups design").[21] This experimental design was chosen for the sake of the exactness of this crucial investigation to avoid possibly confounding "carry-over effects" from one experimental condition to the other.[22] All the *S*s were tested twice: (a) under a "no anchor" condition in order to make a computation of the time-order error constant c in Equation [2] and [3] possible (control session), and (b) 10 to 14 days later under only one "anchor"-condition randomly chosen for a certain *S* (experimental session).

(4) Procedure. The experimental test-situation followed Helson's own as far as possible: his instructions (a) for the 'no anchor' condition, and (b) for the 'anchor' condition (1947),[23] were translated into German with only minor modifications (*e.g.* simplifications of phrases in order not to overstrain the *S*s' intellectual capacity). In addition, a graphical vertical rating-scale containing Helson's nine qualitative categories ranging from "very very light" (*i.e.* "sehr sehr leicht") to "very very heavy" (*i.e.* "sehr sehr schwer") was presented at the screen at eye-level to facilitate the *S*s' judgmental orientation.

In the experiment, *S* sat on a chair at a table, the right elbow resting on a soft trough which was adjustable in respect to *S*s' arm-length. As to the anchor-condition, the same anchor (which was not to be judged) always preceded the series-stimulus, and *S* had merely to judge the second stimulus, namely; the series-stimulus according to the graphical scale. After having given the respective instruction, *E* presented 10 practice runs either together with only the five series weights (a: control session) or with an anchor-weight (b: experimental session).[24] After the 10 practice runs, *E* presented 10 further test-runs which constituted the main experimental phase; the *S*s' judgments at this stage were recorded. Throughout the experiments each series-stimulus in each run was randomly determined.

The time between stimulus-presentations was held as constant as possible by the *E*. The presentation time of a single series-stimulus was, approximately, 1-1.5 sec., the time-interval between series-stimulus and anchor, ca. 1.5 sec., the time interval between two successive pairs of series- and anchor-stimulus ca. 2 sec., the time-interval between two series-stimuli (a: in the control session) ca. 1.5 sec. An experimental session for one S lasted ca. 10 to 15 min. (the presentation of instruction included).

---

[20] This apparatus constructed according to Helson's own experimental requirements has been described in detail in Sarris, *op. cit.*, *Z. exp. angew. Psychol.* 14, 1967, 13.

[21] This type of experimental design is described in A. E. Edwards, *Experimental Design in Psychological Research*, rev. ed., 1960, 117-125; 148-150.

[22] As to the danger of carry-over effects, see Quinn McNemar, *Psychological Statistics*, 3rd ed., 1962, 343 f.; and B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 300 f.

[23] Helson, *op. cit.*, 1947, 7 f.; cited analogously in Helson and Masters, *op. cit.*, 402; see also Sarris, *op. cit.*, *Z. exp. angew. Psychol.* 14, 1967, 20-22.

[24] Helson presented in 1947 only *five* practice runs; however, as Rambo and Johnson have recently shown, there is a strong dependence of *AL*-values from practice. It was decided, therefore, in this study to exclude such confounding effects as far as possible (cf. Helson, *op. cit.*, 1947, 8; W. W. Rambo and E. L. Johnson, Practice effects and the estimation of adaptation-level, this JOURNAL, 77, 1964, 106-110).

RESULTS

Preparatory analysis. Data from 160 *Ss* judging the five series-stimuli 10 times, each of them: (a) under the same control conditions ("no anchor"), and (b) with only one randomly chosen anchor stimulus, were gathered. They yielded altogether 320 different individual judging curves. Since the averaged $J = f$ (S) curves,[25] for each experimental group separately, could be fitted reasonably well by means of the *logarithmic* function, Helson and Himelstein's method for computing the respective 320 empirical *AL*-values was applied.[26] Consequently, since Fechner's law was sufficiently useful as to the present material, the averages for the individual *ALs* were to be represented by *geometric means.* To make appropriate use of the techniques of parametric analyses of variance possible, the logarithmic transformation of the 320 *ALs* was chosen. Further preparatory analysis had the following outcomes:

(1) Control data. By means of an analysis of variance (its theoretical assumptions being met) it was shown that the 10 experimental groups did not differ significantly from each other in respect to their control ("no anchor") log *AL* values ($F = 1.57$, $df = 9/136$, $0.25 > p > 0.10$). This means that according to the purpose of the study the randomization process for gaining the experimental groups had been organized successfully. Therefore, the necessary and minimal requirement for adequate interpretation of the anchor-*AL*-relationship was fulfilled (see below).

The mean empirical *AL*-value (geometrically averaged for the 160 control data) was $AL = 259.05$. This value is significantly larger than Helson's own $AL = 253.9$ computed by respective use of Equation [3] for the five series stimuli ($t = 2.64$, $df > 100$, $p < 0.01$). Therefore, in solving Equation [3] for Helson's constant "time-order error" c by using $AL = 259.05$, the value $c = 0.65$ found in this study is more appropriate (mostly because of the large N) than Helson's $c = 0.75$.[27]

(2) Experimental data. The requirements for homogeneity of variance and normality of the experimental ("anchor") log *ALs* were not met as Bartlett's test and the Chi-square test for goodness of fit indicated (in either case $p < 0.01$) unless Lienert's *"Signum-Wurzel-Transformation"* (signum root transformation) for hyper-excessive distributions was used.[28] If the value to be transformed has the symbol x (= log *AL*), the transformed value x′ is computed according to the following relationship:

---

[25] The symbol *J* means the numerical judgmental value ranging from 1 to 9, S the physical value of the series stimuli.

[26] Helson and Himelstein, *op. cit.*, 631-637.

[27] Cf. Helson, *op. cit.*, 1947, 5; *op. cit.*, 1964, 130 f.

[28] G. A. Lienert, Über die Anwendung von Variablen-Transformationen in der Psychologie, *Biometr. Z.*, 4, 1962, 145-181.

$$x' = \mathrm{sgn}\ (x - \bar{x})\ (1x - \bar{x}1)^{1/2}$$

where $\bar{x}$ is the arithmetic mean for all values to be transformed.[29]

As to the logic of data transformation in general, the reader is referred to Lienert's article.[30]

Main results. (1) In Fig. 1 the theoretical (a) anchor-*AL* relationship according to Helson's Equation [2] is confronted with the empirical (b) anchor-*AL* trend. The empirical *AL*-trend is constituted by the geometric mean values (see above) for the 10 experimental anchor-conditions. The discrepancies between both curves are great, since it appears from Fig. 1 that the experimentally established curve has two inflection-points: the first one occurs at an anchor-value of ca. 75 gm. where *AL* is lowest, and the second one at an anchor-value of ca. 2900 gm. where *AL* is highest. Apparently, the empirical points do not follow a monotonic curvature as it is mathematically implied in Helson's basic Equation [2], but beginning from both inflection-points they seem to approach the control level (parallel to the abscissa at $AL = 259.05$) as the anchor becomes more extreme. Psychologically, this means that with an extreme anchor relative to the stimulus-series the *contrast* effect of the anchor seems to disappear gradually.

To prove the statistical relevance of differences between trend (a) and (b), Lindquist's test for trend differences was applied.[31] This test based upon the principles of analysis of variance allows *independent* investigation of both *pattern* and *vertical displacement* for trend differences. The testing of differences as to the pattern is most relevant in this connection. Since the respective *F*-test yielded a highly significant value ($F = 3.11$, $df = 9/150$, $0.001 < p < 0.01$), it was decided that Helson's anchor—*AL* implication involved in Equation [2] and especially in Equation [3] must be rejected on empirical grounds. By means of further analysis of *variance components*,[32] it was shown that the differences followed a statistically high significant trend which was non-linear, *i.e.* cubic, in character ($F_{\text{nonlin.}} = 3.33$, $df = 8/150$, $p = 0.001$). This additional result gives a sound empirical base for an appropriate *modification* of Helson's basic Equation [2].

(2) In Equation [2], the weighting coefficient *k* has been defined as a *constant*.[33] Now, by use of the numerical anchor values and the respective experimental *AL*s and by solving Equation [2] for *k*, successively, we

---

[29] Lienert, *op. cit.*, 171.

[30] Lienert, *op. cit.*, 178.

[31] E. F. Lindquist, *Design and Analysis in Psychology and Education*, 1953, 344 f.

[32] See Quinn McNemar, *Psychological Statistics*, 3rd ed., 1962, 275-278.

[33] See Helson, *opp. citt.*, 1947, 3 ff.; 1959, 579-584; 1964, 129-133.

find that *k covaried* with size of anchor-stimulus. In Fig. 2, the (asymmetric) parabolic-type relationship between log *C* and *k* is shown. This means psychologically, that for relatively very light and very heavy weights the anchor's effectiveness as to the *Ss*' judgment of series-stimuli is diminishing.

If *k*, as an experimental *variable,* is explicitly taken into account in Equation [2], the empirical *AL*s become predictable as is shown by the cubic curvature in Fig. 1. To fit the empirical $k = f(C)$ relationship, the following mathematical equation was chosen:

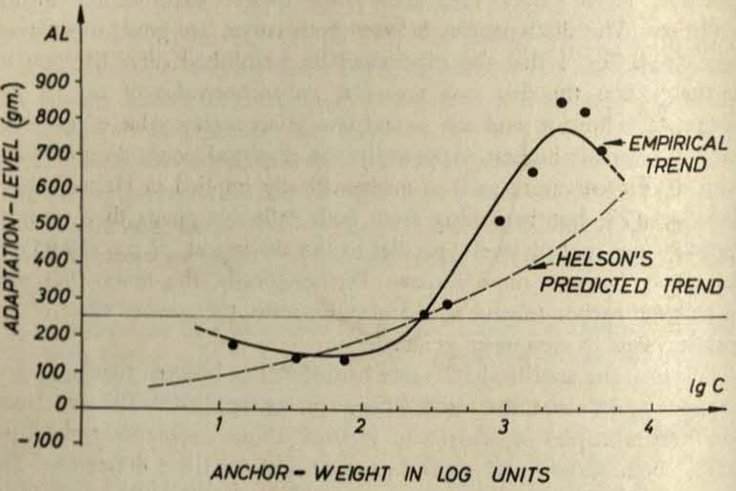


FIG. 1. ADAPTATION-LEVEL AS A FUNCTION OF THE ANCHOR-STIMULUS
Helson's predicted anchor-*AL* trend according to his weighted log mean *AL* formula from 1947 (also 1964) is confronted with the empirical trend. The experimentally established curve appears to be *non-monotonic* in character, and with anchor becoming more extreme it seems to approach the *control AL* level, i.e. the *AL* value for the experiment without anchor (control line parallel to the abscissa at $AL = 259.05$).

$$k' = 1/dC^{e} + f\,C^{g} \qquad [4]$$

where *d, e, f, g* are empirical constants with a random fluctuation. Their values were computed from the present data material as follows (least square method):

$d = 0.01068 \quad e = 0.92983 \quad f = 0.00037 \quad g = 1.01014.$

The application of Equation [4] for fitting k has two important advantages when used in Equation [2]: (a) This equation takes into account the *asymmetric* parabolic relationship between C and k and has, therefore, a satisfactory *predictive* value. (b) This Equation has, moreover, a *theoretical* value, since for $C \to O$ and $C \to \infty$ it follows mathematically that $k' \to \infty$ so that Helson's Equation [2] ("with anchor") changes into Equation [3] ("no anchor"). This mathematical implication is consistent with the psychological interpretation of the anchor's diminishing effectiveness as stated above.

EXPERIMENT II: STEP-INTERVAL-*AL*-RELATIONSHIP

Method: (1) Stimuli. To test the empirical validity of Equation [3a] adequately, it had been necessary to vary the step-interval d in keeping the geometric mean $\bar{S} = 291.4$ used in Equation [3a] constant, at the same time. The following four stimulus-series chosen for this study met these requirements:

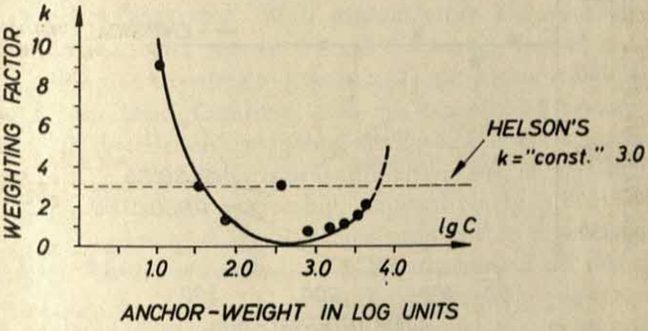


FIG. 2. THE WEIGHTING FACTOR k AS A FUNCTION OF THE ANCHOR-STIMULUS

In contrast to Helson's statement from 1947 (also 1964) that $k =$ const. 3.0 the empirical trend shows k to be a *variable*.

(a) $d =$ *40:* 216.9; 256.9; 296.9; 336.9; 376.9
(b) $d =$ *70:* 168.4; 238.4; 308.4; 378.4; 448.4
(c) $d =$ *120:* 102.4; 222.4; 342.4; 462.4; 582.4
(d) $d =$ *200:* 37.7; 237.7; 437.7; 637.7; 837.7

Since in Helson's quantitative theory the relevance of the d-factor seems to be less important than *e.g.* his anchor-variable, only four experimental conditions were chosen. The cited d-values were irregular in order to test both rather extreme and intermediate d-levels.

The series stimuli were produced by means of different amounts of lead-shot into the same plastic bottles described above for Experiment I.

(2) Apparatus. To present the different stimuli to the *Ss*, the same apparatus mentioned above was used.

(3) Subjects. The *Ss* were 40 right-handed soldiers (not tested in Experiment I), their ages ranging between 19 and 24 yr. As in Experiment I, the soldiers were not

familiar with the problem being investigated and they were randomly assigned to the four different experimental conditions thus 10 *S*s forming an experimental group ("randomized groups design"). All the *S*s were tested only once; namely, under the respective step-interval condition.

(4) Procedure. The experimental procedure used was exactly the same as that for the control ("no anchor") condition in Experiment I. Again, an experimental session for one S lasted ca. 10 to 15 min.

Results. Data from 40 *S*s judging the respective 5 series-stimuli 10 times, each in random order under only one of the 4 step-interval conditions, were analyzed and the *AL*s were computed according to Helson and Himelstein's method mentioned above. With regard to Fechner's law, the empirical *AL*-values were logarithmically transformed and, also in this experiment, *geometric means* were calculated. The locus of empirical *AL*s with varying step-interval *d* ranging from 40 to 200 gm. is shown in Fig. 3 from which it appears that size of step-interval has no important effect on *AL*, if any. The empirical points, apparently best fitted by a horizontal line parallel to the abscissa, are confronted with Helson's own prediction (broken line) according to Equation [3a]. Since a parametric trend analysis[34] resulted in no significant *F*-value either for the linear *variance component* ($F < 1.0$) or for the non-linear one ($F = 1.14$,

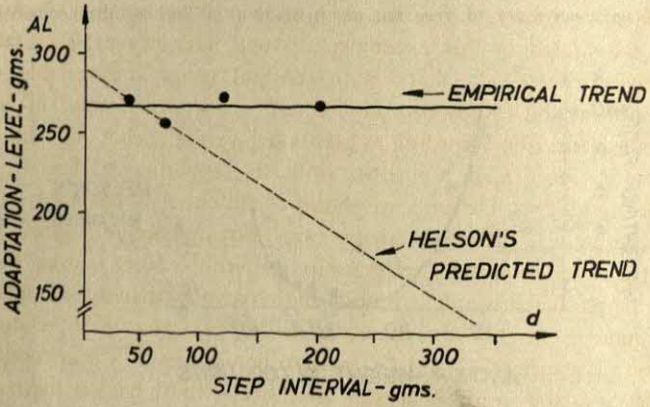


FIG. 3. THE FUNCTIONAL RELATIONSHIP BETWEEN STEP-INTERVAL AND ADAPTATION-LEVEL

Helson's predicted step-interval *AL*-trend according to his weighted log mean *AL*-formula from 1947 (also 1964) is confronted with the empirical results. Since the variation of the step-interval *d* resulted in no statistically significant shift in *AL*, the mean experimental *AL*-values are well represented by a horizontal trend parallel to the abscissa at the general mean $AL = 267.4$.

---

[34] See McNemar, *op. cit.*, 275-278.

$df = 2/36$, $p > 0.25$), it was decided on empirical grounds that the step-interval d was an experimentally *irrelevant* variable as to shifts in *AL* and had been falsely included by Helson in his formulae. As consequence, instead of writing "$c{\cdot}d$" $= 0.65 \cdot 50 = 32.5$ in Equation [2] and [3] it is possible to use simply an empirical constant $s = 32.5$, which, presumably, cannot be further *theoretically* interpreted at the present experimental-scientific stage.

Discussion and conclusions. Helson's mathematical prediction from his weighted log mean *AL*-formula as to his own lifted weights situation has been *invalidated* in this experimental study. Helson's explicit statement of 1947 (also 1964) that "With constant series and variable *C*, the *AL*-varies as the fourth root of *C* . . ."[35] is wrong. Instead of being a monotonic function according to Equation [2], the anchor-*AL* relationship follows a *cubic* trend. Consistent with this experimental finding is the further result that Helson's weighting coefficient k is not a constant but a (asymmetric) parabolic variable with varying size of anchor. Consequently as to *anchor-AL* relationship, empirical reality is even for Helson's simple lifted weights situation more complex than accommodated by *AL*-formula. On the other hand, as to the *step-interval-AL* relationship, Helson's Equations [2] and [3] seem to be unnecessarily over-complicated by the '*d*-factor' since it appears that size-interval has no experimental effect on *AL*.

Without appropriate *modification*, Helson's weighted log mean formula is of questionable value, because, as Stevens has put it against *AL*-theory in another connection, ". . . a quantitative theory that is wrong is not better for being quantitative."[36] Therefore, the following revised formula indicated by the experimental results is proposed instead of Helson's one, as it takes into account in Equation [2] (a) the 'constant' k as a variable v, and (b) the '*d*-factor' $c{\cdot}d$ as a simple empirical constant s

$$\left\{ \begin{aligned} \log (A + s) &= \frac{v \log S + \log C}{v + 1} \\ v &= f\,(C) \end{aligned} \right\} \qquad [5]$$

Analogously, in Equation [3] the term $c{\cdot}d$ should be replaced by the empirical constant s.

As to the *variable v*, beyond its predictive (empirical) value, the quantitative revision has three theoretical advantages: (a) The revised formula takes into account former empirical findings as to anchor-effects made independently from *AL*-theory and summarized, *e.g.* by Guilford: "An anchor may be moved so far from the stimulus range that the shifting of the scale reaches a 'breaking point'. In this event the limen falls back, but not completely."[37] (b) Helson's Equations [2] ("with anchor") and Equation [3] ("no anchor") are linked together upon a *mathematically* sound base and not only upon psychological reasoning or even speculation,[38] because, theoretically, with $v \rightarrow \infty$ Equation [2] changes into Equation [3] as stated above. Helson's own restrictive comment that anchor *in*effectiveness with $C = 0$ (or near zero) ". . . introduces a *discontinuity* in the weighted log mean definition of *AL* . . ."[39] becomes invalid, therefore. (c) The experimentally established *inflection-points* are easily interpretable in the light of theoretical considerations made by Helson himself in another connection, namely:

---

[35] Helson, *op. cit.*, 1947, 5; also *op. cit.*, 1964, 131.
[36] Stevens, *op. cit.*, 637.

> ". . . each stimulus is ordered both as a function of membership in its class and as the result of the adjustment made by the organism to the class as a whole. Thus within classes (houses, dogs, men, automobiles) some members are large, some are medium, and some are small; and what is large in one class may be small compared with the members of other classes."[40]

This means, beyond certain limits (now being predictable from Equation [5], there is a loss of anchor effectiveness, because very extreme anchors relative to the series-stimuli are implicitly experienced as not yet belonging to the possible stimulus-series 'class' which is to be judged by the *S*s.

Though at a first glance the quantitative revision proposed here seems to imply a strong critique against Helson's *theory*, it is more correct to consider the modification of Helson's basic formula as a step consequently taken against Helson's fundamental *mathematical* model. The quantitative revision itself may be even interpreted in honor to *AL*-theory which is a bold theory because of its promising generalizing power. For the statistical falsification of both mathematical implications studied here does not interfere with Helson's purely theoretical (qualitative) concept. In fact, the modification corresponds to Helson's own theoretical suggestion generally made elsewhere: "The simple weighted log mean formula employed by the writer and his coworkers may be modified to take into account cases more complicated than the one with which we have so far dealt."[41] On the other hand, it must be noted, nevertheless, that Helson until now has not been willing to modify his basic formula explicitly concerning the weighting coefficient v in relation to Equation [2].[42] It seems that, for many years, Helson's mathematical model was formulated more precisely than intended for testing and application. For Helson (until 1964) has tried to support his important anchor-*AL*-relationship which is predicted from Equation [2] by chosing only 3 different experimental levels with only 4-6 *S*s for each condition, looking only for graphical evidence.[43] This hesitation over 20 years as to *systematic* tests of the basic *AL*-formula may have been the reason for Helson and Masters' omission to revise the quantitative *AL*-model explicitly according to their own finding that the weighting coefficient for the stimulus-series and the anchor is a variable.[44]

---

[37] Guilford, *op. cit.*, 1954, 313.

[38] Cf. Helson and Masters, *op. cit.*, 400; see also Sarris, *op. cit.*, this JOURNAL, 1967, 304-309.

[39] Helson, *op. cit.*, 1964, 142. (Italics ours).

[40] Helson, *op. cit.*, 1964, 126.

[41] Helson, *op. cit.*, 1959, 582.

[42] See Helson and Masters' article, *op. cit.*, 400-408; also Sarris' critical comments against Helson and Masters' line of reasoning, *op. cit.*, 1967, this JOURNAL, 1967, 304-309.

The presence of inflection-points as established for the anchor-*AL* relationship raises some basic questions to whole *AL*-theory. For example: Are there any safe *a priori* reasons for the fact that in Fig. 1 the upper "tail" of the empirical trend does not seem to reach the horizontal control line in contrast to the lower one, the same being true of Helson and Masters' empirical curve?[45] Will it become possible to construct a *meaningful* mathematical *AL*-model which takes into consideration shifts in *AL* with variation of both anchor *and* stimulus-series, and this not only for lifted-weights but also for other psychophysical or still other materials? What about the generalizing power of quantitative *AL*-theory, if the research worker really has "to investigate locus of *AL*s in each sense modality to determine effects of anchor and other stimuli"[46] and thereby finds that for some psychophysical materials there are inflection-points, but for others none? Is it really true that specific mathematical functions (like for anchor-*AL* curves) are just the same both for physiological and for semantic dimensions? According to Stevens' arguments against *AL*-theory one may answer the latter question perhaps in the negative in stressing the point that

> ". . . we must . . . try to keep separate two important problems—the rules relating to the physiological processes called adaptation and the rules relating to the semantic dispositions that manifest themselves in human judgment. The two problems are related in certain ways and they may sometimes exhibit interesting *interactions*, but they are not necessarily the same thing."[47]

Much more experimental *AL*-work will have to be concerned with these basic problems and will have to supplement possible interpretations of the present experimental results.[48] For these are questions still unsolved which may be asked with full justice in connection with a critical analysis of quantitative *AL*-theory even though their investigation has not been the aim of this study. Here it may be noted in short only, that recent systematic experiments concerning the shift in *AL* with variation of both anchor and stimulus-series for *pitches* give substantial support to the quantitative revision concept suggested in this paper.[49] The main task of the present study has been the test of the basic weighted log mean *AL*-model from 1947

---

[43] See Helson, *op. cit.*, 1947, 8 and 10 f.; 1964, 133 ff.

[44] Helson and Masters, *op. cit.*, *e.g.* 407 f.

[45] Helson and Masters, *op. cit.*, 403 ff.

[46] Helson and Masters, *op. cit.*, 406.

[47] Stevens, *op. cit.*, 634. (Italics ours).

[48] In this connection it is supposed that, among other things, Parducci's important contributions to *AL*-research will have to be taken into account more exhaustively than they have been appreciated by Helson's school hitherto. See *e.g.* Parducci, *op. cit.*, 1963, 1-50.

[49] Experimental research work carried out by the present writer in autumn 1966. The report of this study is in preparation.

(also 1964) in respect to Helson's main psychophysical material (lifted-weights) he has used for his own demonstrations.

SUMMARY

For twenty years, since creating *AL*-theory, Helson has taken the scientific value of his basic weighted geometric average *AL*-model for granted. The *mathematical* implications of his so-called "quantitative" theory have, however, never been sufficiently tested and developed by means of trend-analytical experimentation. As Helson's theory claims many promising results for it and as it may get perhaps great relevance for whole psychology because of its theoretical versatility, a critical experimental analysis of the quantitative model appears to be all the more indispensible. Therefore, in this study, Helson's fundamental *AL*-formula from 1947 (also 1964) was tested in respect to two different mathematical implications: (a) as to the *anchor-AL* relationship (Experiment I) and (b) as to Helson's "*d*-factor," *i.e.* the *step-interval-AL*-relationship (Experiment II).

Both tests were carried out in exact replication of Helson's original experimental procedure with lifted-weights. In Experiment I, 160 *S*s had to judge Helson's series-stimuli of 200, 250, 300, 350, and 400 gm. under 10 anchor conditions the lightest anchor being 12.4 gm., the heaviest being 4500 gm. ("randomized groups design"). In Experiment II, again by use of a randomized groups design, 40 other *S*s had to judge four different series of five weights, the geometric means for these series always being the same and only the step-interval *d* between the series-stimuli being varied. The statistical analysis based upon parametric trend-analytical techniques resulted in a disproof of the quantitative *AL*-model concerning both Helson's anchor-*AL* and step-interval-*AL*-implication: In contrast to Helson's basic *AL*-formulae, the empirical anchor-*AL* relationship is non-monotonic, *i.e. cubic,* in character, and, besides, size of step-interval has no statistically significant effect on *AL*. In respect to anchor-*AL* relationship, further statistical analysis proved Helson's constant weighting coefficient *k* to be a *variable* depending on size of anchor value. By taking into account empirical reality, Helson's basic *AL*-model was respectively revised. It could be demonstrated that the *modification* of the mathematical model has both empirical *and* theoretical advantages.

The presence of inflection-points as a function of both anchor- and stimulus-series raises basic questions to *AL*-theory which are hitherto not sufficiently solved. Therefore, much more critical experimental work has been suggested.

AN ANALYSIS OF CONCEPT-CLUSTERS IN SEMANTIC INTER-CONCEPT SPACE

By JOHN E. HOFMAN, Hebrew University, Haifa College, Israel

The semantic differential is a technique for measuring the meaning of concepts by rating them, "differentiating" them, on a set of bipolar adjectival scales selected to represent an hypothetical multi-dimensional space.[1] One of the features of the techniques is the allocation of the rated concepts to points in semantic space. The distances between points are held to be roughly proportional to whatever psychological affinity between them is mediated by meaning. Thus, when concepts from some domain, say, that of education, are differentiated, they will form configurations composed of meaningfully related conceptual clusters; for example, concepts relevant to one philosophy of education will be relatively closer to one another than they will be to concepts relevant to another philosophy. A cluster can be defined as a grouping of two or more elements more related to one another than they are to elements outside of the grouping.

In their well-known blind analysis of a case of multiple personality, Osgood and Luria demonstrated the usefulness of analyzing clusters in semantic space,[2] but the method does not seem to have been widely applied. The unavailability of suitable quantitative procedures may be responsible for the neglect of a promising analytic tool. Osgood, *et al.*, in discussing the problem of identifying clusters in semantic space, rely mainly on visual inspection but point out that "it would be much more satisfactory to have some method for deriving clusters directly from the *D* matrix."[3]

Several methods of grouping items have been described. Examples are Mc Quitty's linkage analysis;[4] Nunnally's factoring of raw scores;[5] Osgood and Suci's *D* method of factoring;[6] Osgood, Ware, and Morris' application of the centroid method of factoring;[7] Sawrey, Keller, and Conger's method of grouping profiles by distance func-

---

* Received for publication December 20, 1966. The present study is based on a Ph.D. thesis completed at New York University, School of Education, under the sponsorship and with the generous help of Prof. F. N. Kerlinger.

[1] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, *The Measurement of Meaning*, 1957.

[2] C. E. Osgood and Zella Luria, A blind analysis of a case of multiple personality using the semantic differential, *J. abn. soc. Psychol.*, 49, 1954, 579-91.

[3] *Op. cit.*, 102.

[4] L. L. Mc Quitty, Capabilities and improvements of linkage analysis as a clustering method, *Educ. psychol Measmt.*, 24, 1964, 441–56.

[5] L. Nunnally, The analysis of profile, *Psychol. Bull.*, 59, 1962, 311–19.

[6] C. E. Osgood and G. J. Suci, Factor analysis of meaning, *J. exp. Psychol.*, 50, 1955, 325–388.

[7] C. E. Osgood, E. E. Ware, and Charles Morris, Analysis of the connotative

tions,[8] Ward's procedure of hierarchical grouping,[9] or Guttman's non-metric representation.[10]

These methods are actually or potentially useful in assigning *all* variables to *some* cluster (factor, dimension, category, etc.). While this is what we may most often wish to do there are occasions when we want to make statements about individual differences in fact and extent of structurization. One such occasion is the analysis of semantic space. Does any particular person differentiate all of the concepts in some set, or only some of them? Do concepts in any one semantic space form a comprehensive whole, or are they separated into distinct subsystems? Groups of individuals, too, may differ in their manner of ordering concepts into more or less complete and more or less integrated configurations of clusters. A cluster analysis of semantic space should enable one to describe and compare the concept clusters (and non-clusters) of individuals and groups of individuals.

In a recent study of attitudes toward education, the present writer compared cognitive systems formed by the semantic differentiation of educational concepts.[11] An empirical method was described that transformed inter-concept distances (*Ds*), pooled the transformed *Ds* into an overall frequency distribution, and determined a critical distance to define members of clusters. The method would seem to satisfy the condition of deriving clusters almost "directly from the *D* matrix."

Method. The starting point for the present method of cluster-analysis, as already indicated, is *D*, a statistic derived from responses to the semantic differential by the application of the generalized formula of solid geometry.[12] It is expressed as:

$$D_{ij} = \sqrt{\sum (X_i - X_j)^2} = \sqrt{\sum d_{ij}^2}$$

where D_{ij} is the linear distance between the points in the semantic space representing concepts i and j; X_i and X_j are scores of the concepts i and j on the same adjective scale; and d_{ij} is the algebraic difference between the coördinates of i and j on the same k adjective scales. *D*'s can be computed for the distances between each and all concepts, forming a symmetric matrix.

Once a matrix of inter-concept *Ds* is obtained, the next question is: When is a *D* between two concepts small enough relative to other *Ds* so that it defines the concepts it links as a cluster or part of a cluster? Some one distance has to be chosen as a *critical distance* so that it and distances smaller than it will define mem-

---

meaning of a variety of human values as expressed by American college students, *J. abn. soc. Psychol.*, 62, 1961, 62–73.

[8] W. L. Sawrey, J Keller and J. J. Conger, An objective method of grouping profiles by distance functions and its relation to factor analysis *Educ. psychol. Measmt.*, 20, 1960, 651–73.

[9] J. H. Ward and M. E. Hook, Application of an hierarchical grouping procedure, *Educ. psychol. Measmt.*, 23, 1963, 69–81.

[10] A. L. Guttman, The non-metric breakthrough for the behavioral sciences, Proceedings of the Second National Conference on Data Processing, Rehovoth, Israel, January 5–6, 1966.

[11] J. E. Hofman, Attitudes toward education and the meaning of educational concepts, Unpublished Doctoral thesis, School of Education, New York University, 1964.

[12] Osgood, Suci, and Tannenbaum, *op. cit.*, 91.

bers of a cluster. Two problems immediately arise: (1) the distribution of D can be shown to be positively skewed making the application of normal curve statistics inappropriate; and (2) people differ in their response characteristics. These two problems are connected. Since some people tend to use extreme scale positions their D space will exhibit unusually large values even though the relative position of concepts may be as in the space of individuals using more intermediate scale-positions. The transformation of D matrices suitable to the correction of the skew and the manner of selecting a critical distance will be described next. A small contrived example to illustrate these operations is provided in an appendix.

Transformation of D matrix. Ds in each concept matrix are divided one at a time by the square root of the sum of squared distances in that matrix; the sum of the squared matrix elements thus transformed now equals unity (1.0). Or:

$$Dn = \frac{Dq}{\sqrt{\sum_{q=1}^{n} D_q^2}} \quad \text{and} \quad \sum_{q=1}^{n} (Dn_q)^2 = 1.0$$

where Dn is a transformed D, and n is the number of interconcept distances between each and every other concept. The fact that Dns in each individual matrix, when squared, now sum to unity makes matrices comparable and forces the distribution of Dns from pooled matrices into symmetrical form approximating normality. Further operations can be carried out with a normal curve in mind.

Selection of critical distance. The critical distance can usually be chosen by inspection. If concepts cluster at all in at least some of the matrices constituting the distribution, there will be a build-up of 'short' distances somewhere *below* the grand mean. The critical distance should be chosen at the upper limit of the interval exhibiting the most conspicuous bulge, that is, *above the interval in which obtained frequencies most greatly exceed frequencies expected on the basis of a normal distribution.* A researcher operating within the framework of a substantive theory will not be able to depart far from these criteria if he wishes to optimize the formation of hypothesized clusters. Though some flexibility is in order, clusters will become meaningless if the cut-off point is moved up; they will become too few if it is made too conservative. The tendency to maximize meaningful clustering will almost naturally force the student into adherence to the criteria suggested above.

Once a critical distance has been selected it should be applied to *all* matrices of the sample. In the study of attitudes which occasioned the present approach, a Dn of 0.060, corresponding roughly to the lower limit of the 0.20 confidence limits ($z = -1.282$), satisfied the criteria of selection set forth above and maximized hypothesized clusters.

Measures of concept-clustering. Two methods of measuring the extent of clustering in semantic *Dn* space are proposed: by cluster-scores and *Dn* variances. Cluster-scores are simple counts of 'short' distances. Pointing directly to the particular concepts tied into clusters they lend themselves to descriptive purposes. Matrix variances reflect the extent to which inter-concept distances in any one matrix are dispersed. The greater the variance, the greater the differentiation among the cognitive elements of the matrix. Variances have the advantage of taking all the *Dn* information into account.

An application of Dn analysis to the differentiation of educational concepts: As part of the aforementioned study of attitudes toward education, 228 students and professors were asked to rate 13 educational concepts on a seven-step form of the semantic differential containing 12 bipolar adjective pairs. Table I lists concepts and adjective pairs in the order of presentation.

TABLE I

SEMANTIC DIFFERENTIAL OF EDUCATIONAL CONCEPTS

Concepts*	Adjective Pairs
EDUCATION	
CHILDREN'S NEEDS	active—passive
TRADITION	bitter—sweet
CREATIVITY	strict—lenient
SUBJECT MATTER	boring—interesting
PERSONALITY GROWTH	stable—changeable
DISCIPLINE	aloof—responsible
TEACHER	soft—hard
MORALITY	definite—uncertain
SELF EXPRESSION	warm—cold
AUTHORITY	relative—absolute
PUPIL INTEREST	dependent—independent
ME	pleasant—unpleasant

* Each concept is rated on all adjective pairs in turn.

D and *Dn* matrices were generated for each of the 228 *Ss*. The critical distance, as mentioned above, was set at $Dn = 0.060$, from the combined frequency distribution of all 228 *Dn* matrices. To illustrate the application of cluster analysis, the *Dn* matrices of two individuals will be compared (Table II). The two are women professors of education at a large urban university. Both are high scorers on a scale of educational progressivism. Both have highly differentiated semantic spaces. But the manner in which

TABLE II

DN MATRICES OF TWO PROFESSORS OF EDUCATION

A. Guidance Expert: Cluster Score: 18*; Matrix Variance: 0.00243†

Concepts	2	3	4	5	6	7	8	9	10	11	12	13
1. EDUCATION	044	052	060	055	063	042	052	180	131	198	160	088
2. TEACHER		065	058	073	064	023	046	161	104	176	137	073
3. ME			095	053	058	048	042	156	116	173	136	073
4. CHILDREN'S NEEDS				086	068	068	077	200	146	214	177	116
5. CREATIVITY					048	064	068	176	124	196	149	084
6. PERSONALITY GROWTH						157	152	172	122	185	147	088
7. SELF EXPRESSION							035	151	098	165	126	061
8. PUPIL INTEREST								147	101	162	128	068
9. TRADITION									094	044	069	102
10. DISCIPLINE										093	052	069
11. MORALITY											*053*	118
12. AUTHORITY												082
13. SUBJECT MATTER												

B. Early Childhood Specialist: Cluster Score: 21*; Matrix Variance: 0.00313†

1. EDUCATION	047	102	059	047	080	071	098	106	047	071	221	069
2. TEACHER		110	051	044	069	069	093	107	036	064	200	071
3. ME			124	113	135	112	144	116	113	135	262	107
4. CHILDREN'S NEEDS				051	059	047	074	101	057	047	210	*051*
5. CREATIVITY					*059*	064	074	107	136	064	213	062
6. PERSONALITY GROWTH						071	*051*	109	064	*051*	225	064
7. SELF EXPRESSION							071	086	074	057	264	031
8. PUPIL INTEREST								117	096	076	240	064
9. TRADITION									104	096	194	076
10. DISCIPLINE										059	199	071
11. MORALITY											200	059
12. AUTHORITY												200
13. SUBJECT MATTER												

Note: All decimals have been omitted. (Ex. 044 instead of 0.044)

* Mean cluster score for all 228 matrices: 6

† Mean matrix-variance for all 228 matrices: 0.00125

they structured their concepts revealed important differences in meanings associated with concepts and, perhaps, in cognitive functioning.

The distribution of 'short' distances, underlined in Table II, in the case of the first professor, a guidance specialist, indicates a separation of concepts into two distinct clusters such that distances within the clusters are shorter than between them. One cluster contains concepts usually associated with a 'progressive' viewpoint in education: CHILDREN'S NEEDS, CREATIVITY, PERSONALITY GROWTH, SELF EXPRESSION, and PUPIL INTEREST. The second cluster includes "traditionalist" concepts: TRADITION, MORALITY, AUTHORITY, and DISCIPLINE. The *S*'s progressivism is indicated by the fact that ME, EDUCATION, and TEACHER form part of the 'progressive' cluster. The only concept among the 13 to remain aloof is SUBJECT MATTER; it does not seem to fit into her system, possibly because it is meaningless to her, or irrelevant.

The other professor, a friend of the first and an expert in early childhood, ordered the same concepts quite differently. Knowing her to be a 'progressive', we note that she appears to have interpreted them with reference to one dominant viewpoint. Rather than to leave antithetical concepts apart, she has incorporated them into one comprehensive system. DISCIPLINE, for example, which has 'short' connections with CHILDREN'S NEEDS, MORALITY, and EDUCATION may have a meaning to her quite different from that usually associated with it, perhaps SELF-DISCIPLINE. The most central concept in her system, judging by the number of its tie-ups with other concepts, is CHILDREN'S NEEDS which is not surprising in an early childhood specialist. One of the concepts with no tie-ups at all is AUTHORITY which she keeps far from her system. The fact that ME is aloof from her clusters, too, makes one wish to ask further questions.

Clearly, to the two *S*s of our comparison, the same concepts have different connections or meanings. More crucially, one notes different styles of structuring: one professor sharpens distinctions assigning the various concepts clearly delineated positions in the semantic space; the other scans the concepts as a totality and assimilates them to her point of view.

Whenever a semantic differential is composed of a set of concepts representing some universe of discourse and administered to a sample of individuals, comparisons among semantic spaces of individuals and criterion groups would seem to benefit by cluster analysis. People having 'similar' measurable attitudes tend to assign different meanings to the elements of such attitudes, and they tend to differ in the ways they incorporate part or all of these elements into one system.

SUMMARY

An empirical method of identifying and describing clusters formed in the hypothetical space generated by the semantic differential technique has been explained and an example of its use reviewed. The method includes the generation of a *D* matrix, its transformation to a *Dn* matrix for purposes of comparison and statistical treatment, and the selection of a critical distance by the application of normal curve statistics. The usefulness of the methods has been illustrated by the formation of clusters coalescing around the educational dimensions of "progressivism" and "traditionalism" in the semantic spaces of two professors of education.

APPENDIX

The several steps leading from a semantic differential score matrix to measures of concept clustering are illustrated in the following contrived example:

(1) Scoring. After individuals check one of seven steps on each bipolar adjectival scale for each concept they are assigned corresponding scores from 1 to 7. Thus an adjective-pair-by-concept matrix is generated for each individual. If desired, individual score matrices can be summed and averaged as will be shown.

Assuming three individuals, Tom, Dick, and Harry, three concepts, A, B, and C, and three adjective pairs, 'good-bad,' 'fast-slow,' and 'strong-weak'; the following score matrices may be imagined.

Scales	Tom's Concepts			Dick's Concepts			Harry's Concepts			Opposites
	A	B	C	A	B	C	A	B	C	
Good	1	6	1	3	5	3	3	4	3	Bad
Fast	1	7	7	3	5	5	3	5	3	Slow
Strong	7	1	1	5	2	3	5	4	3	Weak

It will be noted that Tom differs from both Dick and Harry in "response style" as he prefers extreme scores to their more intermediate ones. It may also be recognized that when a mental correction for extremeness of scores is made Tom and Dick appear to be more similar in the way they rate these concepts than either is to Harry. In fact, Harry appears to differentiate these concepts very little. It is the purpose of cluster analysis to correct for response style and to bring out both inter-person differences and intra-person idiosyncrasies in clustering.

(2) D matrices. *D* matrices are computed by the application of the generalized formula of solid geometry (see text). To take Tom's scores matrix as an example one obtains: *D* between *A* and *B:*

$$\sqrt{(1-6)^2+(1-7)^2+(7-1)^2}=\sqrt{97}=9.85$$

Similarly *D*s between *A* and *C*, and *B* and *C* can be found to be $\sqrt{72}$ (8.49), and $\sqrt{25}$ (5.00), respectively. Computing *D*s for the three matrices in the same manner, one obtains the following matrices: (D^2s are indicated in the upper right mirror half; (*D*s in the lower left:)

	Tom			Dick			Harry			
	A	B	C	A	B	C	A	B	C	
A	—	$\sqrt{97}$	$\sqrt{72}$	—	$\sqrt{17}$	$\sqrt{8}$	—	$\sqrt{6}$	$\sqrt{4}$	A
B	9.85	—	$\sqrt{25}$	4.12	—	$\sqrt{5}$	2.45	—	$\sqrt{6}$	B
C	8.49	5.00	—	2.83	2.24	—	2.00	2.45	—	C

Differences in what appears to be response style (between Tom and the others) and concept clustering (between Harry and the others) are here indicated in their relative magnitude and distribution of *D*'s. Values are not directly comparable, and there is no objective way of deciding which *D*'s determine clusters. The next few steps are to remedy these shortcomings.

(3) Transforming D matrices. *D*'s are transformed into *Dn*'s by first adding the squared *D*'s of a matrix, extracting the square root from the sum, and then dividing each *D* in turn by this value (see text for formula). The computation is illustrated for Tom's matrix. Computational steps are shown in the upper right mirror half; *Dn*'s, in the lower left:

Tom's Dn Matrix

Concepts

	A	B	C
A	—	$\frac{9.85}{\sqrt{97+72+25}}$	$\frac{8.49}{\sqrt{97+72+25}}$
B	.71	—	$\frac{5.00}{\sqrt{97+72+25}}$
C	.61	.36	—

As a check on computations, *Dn*s, when squared, sum to unity, within rounding errors: $(.71)^2 + (.61)^2 + (.36)^2 = 1.01$. The other *D* matrices are transformed in the same way. (Computer programs can easily be developed). When placed side by side, the three *Dn* half matrices look as follows:

	Tom			Dick			Harry		
	A	B	C	A	B	C	A	B	C
A		.71	.61		.75	.52		.61	.50
B			.36			.41			.61

Tom and Dick are now seen to differentiate in similar manner, while Harry differs from both.

(4) Determination of critical distance. The distribution of frequencies in the three *Dn* matrices is represented next. For purposes of comparison the frequency distribution of *D*s (before transformation) is shown next to it.

Dn	*f*	*D*	*f*
.30–.49	2	2–3.99	5
.50–.69	5	4–5.99	2
.70–.89	2	6–7.99	0
	–	8–9.99	2
	9		–
			9

It will be noted that the distribution of *D*s is positively skewed, while the *Dn* matrix is symmetrical. In fact, no clustering is apparent in the nine *Dn*'s of the three pooled matrices, there being no "bulge" in the one interval below the mean. If we wish to assume clustering, anyway, we may choose 0.49 as a critical distance since this is the upper limit of the one interval below the mean. Using a criterion of the lower limit of the 0.20 confidence interval as an alternative we arrive at a value of 0.41, or 0.56 — (1.282) (0.12) = 0.41, the mean of the distribution being 0.56 and the standard deviation, 0.12.

Either standard, 0.49 or 0.41, leads to the same conclusion in this case. The *Dn*s between concepts *C* and *D* in the matrices of Tom and Dick are shorter than either standard, 0.49 or 0.41, and therefore define the two concepts to form clusters in those semantic spaces. None of Harry's concepts cluster by these standards.

(5) Measures of clustering. Cluster scores, *i.e.* the number of *Dn*'s as small or smaller than the critical distance, and matrix variances, that is, the variance of *Dn*s are as follows:

	Cluster Score	*Dn* Variance
Tom	1	.0213
Dick	1	.0212
Harry	0	.0013, least differentiated

(6) Group matrices. The three score matrices can be added, averaged and treated like an individual matrix. Score, *D*, and *Dn* matrices for the three-person group would look like this:

	Group Score Matrix				Group D Matrix				Group *Dn* Matrix		
	A	B	C			A	B	C	A	B	C
Good	2.3*	5.0	2.3	Bad	A	—	$\sqrt{30.4}$	$\sqrt{18.9}$		.73‡	.58
Fast	2.3	5.7	5.0	Slow	B	5.5†	—	$\sqrt{7.8}$			.37
Strong	5.7	2.3	2.3	Weak	C	4.3	2.8	—			

* 2.3 is the mean of the scores which Tom, Dick and Harry severally gave to Concept *A* on the "good-bad" scale.

† $\sqrt{(2.3-5.0)^2 + (2.3-5.7)^2 + (5.7-2.3)^2} = \sqrt{30.4} = 5.5$

‡ $\dfrac{5.51}{\sqrt{30.4+18.9+7.8}} = .73$

When either of the critical distances deemed appropriate are applied to the group *Dn* matrix, the concepts *B-C* are revealed as forming a cluster. The cluster score of the group matrix is 1.0; the matrix variance (of 0.73, 0.58, and 0.37) is 0.0212.

ASSOCIATION AND INTEGRATION IN SERIAL LEARNING

By ZAKHOUR I. YOUSSEF, Eastern Michigan University

Until recently, many S-R learning theorists have considered serial learning as a process of establishing associative connections among adjacent items of the serial list.[1] This position has recently been attacked on the basis of work by Primoff and Young.

Primoff[2] and Young[3] had *Ss* learn a serial list, followed by a paired-associate list consisting of adjacent items in the original serial list. These studies failed to show positive transfer, and were interpreted as suggesting that associations are not formed between adjacent items in serial learning.

Ebenholtz,[4] Winnick and Dornbush,[5] and Slamecka[6] have suggested that association between adjacent items plays little or no part in serial learning, and that the primary determinants of correct responses are cues of serial-position and acquired availability of the responses.

While the evidence presented by Ebenholtz indicating the relevance of positional cues is strong,[7] the present author believes that it is premature to discount interimistical associations. It is the writer's position that serial learning involves two processes. The first is the process of establishing associative connections between adjacent items in serial learning and be-

---

* Received for publication October 27, 1966. This article is based on a dissertation submitted to the Psychology Department of Wayne State University in partial fulfillment of the requirements for the Ph.D. degree in Psychology. This research was supported by Grant G 21476 from the National Science Foundation and directed by Dr. Eli Saltz.

[1] C. L. Hull, The conflicting psychologies of learning—a way out, *Psychol. Rev.*, 42, 1935, 491-516; J. A. McGeoch, The direction and extent of intra-serial associations at recall, this JOURNAL, 48, 1936, 221-245; E. J. Gibson, A systematic application of the concepts of generalization and differentiation to verbal learning, *Psychol. Rev.*, 47, 1940, 196-299; B. R. Bugelski, A remote association explanation of the relative difficulty of learning nonsense syllables in a serial list, *J. exp. Psychol.*, 40, 1950, 336-348.

[2] Ernest Primoff, Backward and forward associations as an organizing act in serial and in paired-associate learning, *J. Psychol.*, 5, 1938, 374-395.

[3] R. K. Young, Paired-associate learning when the same items occur as stimuli and responses, *J. exp. Psychol.*, 61, 1961, 315-318; Tests of three hypotheses about the effective stimulus in serial learning, *ibid.*, 63, 1962, 307-313.

[4] S. M. Ebenholtz, Position mediated transfer between serial learning and a spatial discrimination task, *J. exp. Psychol.*, 65, 1963, 603-608.

[5] W. A. Winnick and R. L. Dornbush, Role of position cues in serial rote learning, *J. exp. Psychol.*, 66, 1963, 419-421.

[6] N. J. Slamecka, Inquiry into the doctrine of remote associations, *Psychol. Rev.*, 71, 1964, 61-76.

[7] Ebenholtz, *op. cit.*, 603-608.

tween the items and their serial position. This is followed by another process; namely, integrating the list into one response.

Language behavior illustrates what is meant in this paper by the integrative process. It is commonly-accepted knowledge that letters become integrated into unified words and thus a combination of several letters becomes one single response. Murdock found little difference in the retention of three-consonant syllables and three words but a great difference between the retention of a three-letter meaningful word and either three consonants or three words.[8] As serial learning increases, the items in the list become integrated into one single response in the same fashion a certain combination of letters becomes integrated into one single meaningful word; items lose their stimulus-properties relative to evoking the subsequent item in the list. The writer hypothesizes that no transfer was obtained in the Primoff and Young studies because serial learning was extended to include the integrative phase which would mitigate against any transfer to the paired-associate learning. One would predict negative or no transfer in these studies if it is assumed that the serial list becomes integrated into one structural unit, and any break-up of that integrated whole into segments—in this case, pairs—would prevent any positive transfer, and it may even induce negative transfer.

Thus, it is hypothesized that if different groups of *S*s learn a serial list to different criteria, and are then tested on a paired-associate list in which the pairs are adjacent items from the serial list, transfer will be a convex curvilinear function of the number of trials presented on the serial list. The function is predicted to rise first, because the associative bonds formed during serial learning can be utilized in the paired-associates task, but with more serial learning the function drops since the serial list becomes integrated, and the transfer effect due to such associative connections approaches zero.

It is also hypothesized that when the pairs of the paired-associates are selected randomly from the serial list (*i.e.* nonadjacent items in the serial list constitute a given S-R pair in the paired-associates list), then the transfer from serial learning to paired-associate learning is also a curvilinear function, but this time, a concave curvilinear function. The function decreases first, because the associative bonds formed during serial learning are a source of interference in the paired-associate task, but with more serial learning the function turns upward since the list becomes integrated, and the transfer effect due to these associative bonds approaches zero.

---

[8] B. B. Murdock, Jr., The retention of individual items, *J. exp. Psychol.*, 62, 1961, 618-625.

Method: (1) Subjects. Two hundred *Ss* from classes in introductory psychology were tested. They were randomly assigned to the groups of 20 each.

(2) Materials. List 1 was a serial list of 20 meaningful and highly frequent and low similar nouns selected from Thorndike and Lorge.[9] List 2 consisted of exactly the same items as List 1, except that the words appeared in a different serial order. List 3 was a paired-associate list, the 10 pairs of which were constructed from adjacent items in List 2. In this paired-associate task, no item appeared as both stimulus and response. The pairs appeared in four different orders. Table I indicates the items in these lists.

TABLE I

MATERIAL PRESENTED FOR FREE-RECALL, SERIAL, AND PAIRED-ASSOCIATE LEARNING

List I	List II	List III
(Serial Task for Nonadjacent Groups)	(Serial Task for Adjacent Groups)	(Paired-Associates Task for all Ss)
COUNTRY	COUNTRY	COUNTRY-HUSBAND
LETTER	HUSBAND	STREET-PROMISE
GARDEN	STREET	GARDEN-WINDOW
RETURN	PROMISE	PROBLEM-OFFICE
PROBLEM	GARDEN	FAMILY-REPLY
ORDER	WINDOW	MATTER-RETURN
MATTER	PROBLEM	JOURNAL-DINNER
PROMISE	OFFICE	STORY-ORDER
JOURNAL	FAMILY	TABLE-LETTER
WINDOW	REPLY	QUESTION-BUILDING
HUSBAND	MATTER	
DINNER	RETURN	
QUESTION	JOURNAL	
STORY	DINNER	
FAMILY	STORY	
STREET	ORDER	
TABLE	TABLE	
BUILDING	LETTER	
REPLY	QUESTION	
OFFICE	BUILDING	

(3) Procedure. The *Ss* were presented with one of the two serial lists (List 1 or List 2) for 7, 14, 21, or 28 trials. Forty of the *Ss* did not receive any serial learning. Half of these were randomly treated as an "Adjacent (*AO*) subgroup" in the analysis of variance; the other half were treated as a "Nonadjacent (*BO*) subgroup" in the analysis. This division was done only for the sake of balance in a two-way analysis of variance. Table II indicates the treatment each of the subgroups received on serial learning and the name given to each subgroup. Finally, all the *Ss* were given 15 trials on the paired-associates list or until one perfect criterion, whichever was achieved first.

---

[9] E. L. Thorndike, and Irving Lorge, *The teacher's word book of 30,000 words,* 1944.

The serial lists were run on a Stowe memory drum at a 2-sec. rate with an 8-sec. intertrial interval. The paired-associate list was run on the same drum with a 2-sec. intratrial interval and a 6-sec. intertrial interval. The paired-associate learning was performed by presenting the stimuli and their respective responses on one trial and the stimuli alone of the following trial. There were four different random orders on the paired-associate learning.

This design, summarized in Table II, has the advantage that each subgroup of BO (the Nonadjacent groups) acts also as a control group for stimulus and response familiarization as well as for warm-up effect in each subgroup of Adjacent (AO) groups.

Results. On the basis of the position advanced by the writer, it was predicted that in learning a paired-associate list, the effect of associations between adjacent items from a previous serial list would be evident only after moderate degrees of serial learning. Small degrees of serial learning will be inadequate to produce associations between adjacent items; by the time *S* reaches criterion on the serial list, the list will have been well enough integrated into a single response-unit that the stimulus-properties of individual items will have been lost. Table III presents the relevant data on number of correct responses in paired-associate learning as a function of serial-learning trials.

TABLE II

TREATMENT RECIEVED BY EACH SUBGROUP ON SERIAL LEARNING

	Trials				
List	0	7	14	21	28
(1) Non-Adjacent	BO	B_7	B_{14}	B_{21}	B_{28}
(2) Adjacent	AO	A_7	A_{14}	A_{21}	A_{28}

TABLE III

MEAN NUMBER CORRECT DURING THE 15 TRIALS OF PAIRED-ASSOCIATE LEARNING AS A FUNCTION OF NUMBER OF PRIOR SERIAL LEARNING TRIALS FOR ADJACENT AND NON-ADJACENT SERIAL PAIRS

	Trials					
List	0	7	14	21	28	Grand mean
Adjacent	115.40	112.40	119.10	115.50	110.55	114.59
Non-adjacent	111.20	102.40	95.95	81.45	114.40	101.08

A 2 × 5 analysis of variance shows that the adjacent items from the serial list were better learned than the nonadjacent, $F\ (1/190) = 13.35$, significant beyond the 0.001 level. Also consistent with the prediction, the effect of adjacent vs. nonadjacent pairs interacted with number of previous serial learning trials, F (4/190 = 3.34, significant at beyond the 0.02 level.

A more exact test of the hypothesized relationships was undertaken by means of orthogonal polynomial analysis. The orthogonal polynomial analysis over the adjacent (*AO*) groups does not support the hypothesis that the performance of the adjacent groups on the paired-associate task is a convex curvilinear function of the number of trials on serial learning. The quadratic component was far from being significant, F (1/190) = 0.42. Since none of the orthogonal polynomial comparisons on the adjacent (*AO*) groups was significant, it is concluded that there are no significant differences among these groups.

TABLE IV

ORTHOGONAL POLYNOMIAL COMPARISON ON THE TOTAL NUMBER OF CORRECT RESPONSES ON THE PAIRED-ASSOCIATE TASK

Source	*df*	*SS*	*MS*	*F*	*P*
Adjacent Groups	4	858.85	214.71	0.31	
Linear	(1)	87.12	87.12	0.13	
Quadratic	(1)	288.05	288.05	0.42	
Cubic	(1)	244.21	244.21	0.36	
Quartic	(1)	239.45	239.45	0.35	
Nonadjacent Groups	4	13,863.92	3465.98	5.07	0.01
Linear	(1)	423.41	423.41	0.62	
Quadratic	(1)	8132.43	8132.43	11.90	0.001
Cubic	(1)	4068.02	4068.02	5.95	0.05
Quartic	(1)	1240.06	1240.06	1.81	
Interaction	4	9151.00	2287.75	3.35	0.05
Linear	(1)	63.21	63.21	0.09	
Quadratic	(1)	5740.80	5740.90	8.40	0.01
Cubic	(1)	3152.83	3152.83	4.61	0.05
Quartic	(1)	195.04	195.04	0.29	
Within	190	129,863.00	683.49	—	—

The orthogonal polynomial analysis over the nonadjacent (*BO*) groups, however, supports the hypothesis that the performance of the nonadjacent groups on the paired-associate task is a concave curvilinear function of the number of trials on the serial task. The quadratic component is significant beyond the 0.001 level of confidence, F (1/190) = 11.90. This analysis also supports the implication that the quadratic component of the interaction between trials and adjacent vs. nonadjacent pairs is significant beyond the 0.01 level.

A summary of the orthogonal polynomial comparison is presented in Table IV.

Young found that his adjacent (*AO*) group was superior to the control group only on the early trials of paired-associate learning;[10] it is suspected that the paired-associate task was too easy to differentiate among the adjacent groups over 15 trials of learning. The same set of analyses were conducted on the sum of correct responses of the first and second trials on the paired-associate list.

The analyses on the first and second trials yield similar results to those conducted on the total number of correct responses summed over 15 trials of paired-associate learning. These analyses, however, indicate that the effects are more pronounced at early stages of learning and they lend some support to our notion that the paired-associate list is too easy to distinguish among the adjacent (*AO*) groups. The quadratic component of the adjacent groups, although still not significant at the 0.10 level of confidence, has an F ratio of 1.94, $d.f. = 1/190$ much higher than that based on total number of correct responses for which F is only 0.42, $d.f. = 1/190$.

Four t-tests were conducted on the total number of correct anticipations on the serial task, each t for testing the difference between one of the *AO* (Adjacent) groups and a *BO* (Nonadjacent) group receiving the same number of trials on serial learning. None of these ts approached the 0.10 level of significance. This would indicate that the Experimental groups were of equivalent ability in serial learning and that the upswing for the nonadjacent condition is not the result of superior *S*s in the B 28 condition. The means and standard deviations on serial learning are shown in Table V.

TABLE V

MEANS AND *SD* OF TOTAL NUMBER OF CORRECT ANTICIPATIONS ON SERIAL LEARNING

Adjacent Groups (AO)	X	SD	Nonadjacent Groups (BO)	X	SD
A_7	48.15	18.23	B_7	45.75	20.61
A_{14}	120.64	51.12	B_{14}	106.65	42.83
A_{21}	225.10	72.04	B_{21}	224.32	62.14
A_{28}	362.16	75.10	B_{28}	369.10	73.16

Discussion. The present results tend to support the conception that two processes are operating in serial learning and that the process of establishing associative connections precedes the process of integrating the list into one response unit. The finding that the paired-associate learning of the Adjacent (*AO*) groups was significantly superior to that of the Nonadjacent (*BO*) groups supports the assumption that associative connections are established between adjacent items in serial learning. The significance of the quadratic polynomial comparison over the Nonadjacent (*BO*) groups supports the integrative process assumed by the present writer. It is obvious that the integrative process follows the process of establishing associative connections between two adjacent items. The paired-associate performance of the Nonadjacent groups begins to rise only after 21 trials of serial learning, indicating the effect of integration.

---

[10] Young, *op. cit., J. exp. Psychol,* 63, 1962, 307-313.

One of the hypotheses deduced from the conceptualization, however, is not so well supported. This hypothesis states that the paired-associate learning of the Adjacent groups will be a convex curvilinear function of the number of trials administered on serial learning. While the results tend to be in the predicted direction, they do not approach any degree of significance except when the analysis is conducted over the first two trials of paired-associative learning. Two factors may be operating against positive transfer among the Adjacent groups: First, the paired-associate list may not be sufficiently difficult to show significant positive transfer over 15 trials of learning (*i.e.* a ceiling effect may be operating). Obtaining a larger quadratic effect over the Adjacent (*AO*) groups and augmenting the significance level of our predictions when the analysis is based on the first two trials lends support to this possibility.

The second factor is that serial position cues which are a source of negative transfer may be mitigating against the positive transfer produced by the associative connections established during serial learning. That is to say, responses become associated with their respective serial positions during serial learning. Since the items in the paired-associate learning occupied positions which were different from those they had in serial learning, some negative transfer is expected. This means that serial position effect must work *against* our prediction of convex curvilinearity for the Adjacent groups, while it must work *with* our prediction of concave curvilinearity for the Nonadjacent group. The support for serial position cues comes from Young.[11] He found that items retaining their serial positions were better retained than items which did not. Ebenholtz found that serial learning is possible by associating the items with their respective serial positions, without forming associative connections among the items themselves.[12] The serial-position effect, however, cannot account for all of our results. The fact that the Adjacent groups were superior to the Nonadjacent groups on the paired-associate learning indicates that associative connections among adjacent items have been established during serial learning. This finding that serial learning is a joint function of associative connections established among adjacent items in the serial list as well as of associative connections among the items and their respective serial positions is contrary to the assertions of Slamecka.[13] He rejects the idea that associative connections among items are formed during serial learning, stating that serial learning involves only the learning of the items, *per se,* and then associating these items with their respective serial position. The adjacent groups in this experiment had *no* advantage over the Nonadjacent groups in either learning the items *per se,* or in serial position effect. Therefore, one must attribute the superiority of the Adjacent groups on the paired-associate task only to the associative connections among adjacent items formed during serial learning.

---

[11] *Idem.*

[12] Ebenholtz, *op. cit.*, 603-608.

After completion of this research, a paper by Battig, Brown, and Schild came to the writer's attention.[14] These authors find that transfer of adjacent associations to a second serial list is maximal at an intermediate degree of original serial learning, lending support to the integrative hypothesis proposed by this writer.

Summary. This writer assumes that serial learning consists of two processes. The first is the process of establishing associative connections between adjacent items. This is followed by the integrative process which transforms the whole serial list to one response unit.

It is deduced from these assumptions that transfer from serial learning to the paired-associates learning of the same list is a curvilinear function of the number of trials on serial learning. Transfer would presumably be positive if the pairs were selected from adjacent items in the serial list and negative if they were selected from nonadjacent items. Such transfer —negative or positive—is an increasing function of serial learning, up to a point, after which it will become a decreasing function of serial learning due to the mitigating effects of the integrative processes. The results of this study support the writer's two-process conception of serial learning.

---

[13] Slamecka, *op. cit.,* 61-76.

[14] W. F. Battig, S. C. Brown, and M. E. Schild, Serial position and sequential associations in serial learning, *J. exp. Psychol.,* 67, 1964, 449-457.

DISCRIMINATION-REVERSAL, EXTINCTION, AND ACQUISITION AFTER DIFFERENT AMOUNTS OF TRAINING

By ISABEL M. BIRNBAUM, University of California, Irvine

The overlearning-reversal effect (*ORE*) may, in principle, be accounted for in terms of a difference between the criterion-group (*C*) and the overtrained group (*OT*) either in rate of extinction of response to the previously positive stimulus (PS^+)[1] or in rate of acquisition of response to the previously negative stimulus (PS^-).[2] According to the extinction-hypothesis, the additional reinforcements given the *OT* group lead to a substantial reduction in resistance to extinction of response to PS^+; support for the hypothesis has appeared in the latency-measures of an experiment on brightness-discrimination with single stimulus presentation.[3] According to the acquisition-hypothesis, overtraining reduces the tendency to avoid the negative stimulus, thus facilitating acquistion of response to PS^- when it is reinforced in reversal; consistent with this hypothesis is the finding that forced trials to the negative stimulus during overtraining (which presumably maintain avoidance of the negative stimulus) eliminate the *ORE*.[4] It should be noted, however, that this outcome also may be accounted for in terms of the extinction-hypothesis: resistance to extinction of response to PS^+ is greater following training with both S^+ and S^- than following training with S^+ alone.[5]

The present experiment, intended to yield the *ORE* in brightness-dis-

---

* Received for publication August 17, 1966. This report is based upon a dissertation submitted to the Graduate Division of the University of California, Berkeley, in partial fulfillment of the requirements for the Ph.D. degree. The research was supported by a USPHS predoctoral fellowship. The author wishes to thank Donald A. Riley for his valuable advice and assistance throughout the course of the study. The helpful suggestions of Leo J. Postman and Rheem F. Jarrett are gratefully acknowledged.

[1] E. J. Capaldi and H. W. Stevenson, Response reversal following different amounts of training, *J. comp. physiol. Psychol.*, 50, 1957, 195–198.

[2] M. R. D'Amato and Harry Jagoda, Effects of extinction trials on discrimination reversal, *J. exp. Psychol.*, 59, 1960, 254–260; Analysis of the role of overlearning in discrimination reversal, *J. exp. Psychol.*, 61, 1961, 45–50.

[3] David Birch, J. R. Ison, and S. E. Sperling, Reversal learning under single stimulus presentation, *J. exp. Psychol.*, 60, 1960, 36–40.

[4] D'Amato and Jagoda, *op. cit.*, 1961.

[5] M. R. D'Amato, Donald Schiff, and Harry Jagoda, Resistance to extinction after varying amounts of discriminative or nondiscriminative instrumental training, *J. exp. Psychol.*, 64, 1962, 526–532; H. M. Jenkins, The effect of discrimination training on extinction, *J. exp. Psychol.*, 61, 1961, 111–121.

crimination, was designed to assess independently the rate of extinction of response to PS^+ and the rate of acquisition of response to PS^-.

Method: (1) Subjects. The *Ss* were 120 male, albino rats, 74-94 da. old.

(2) Apparatus. A Y-maze with 13-in. alleys was used for discriminative training and subsequent testing. Bulbs were mounted above milk glass at the top of each alley, and degree of illumination served as the positive or negative cue. The illumination was 0.05 ft.-c. in a dark alley (D), and 22.2 ft.-c. in a bright alley (B). Starting-time and running-time were measured separately by electric timers (T_1 and T_2). Raising the start-box door activated T_1. When S entered the alley, T_1 was stopped and T_2 was activated. T_2 was stopped when S entered the goal-box.

(3) Design. The *Ss* were run to criterion (Groups *C* and *AC*) on original learning *OL*, or were given 200 or 400 trials of overtraining. At the beginning of training, the *Ss* in Groups *C*, 200, and 400 were 74 days old. Those in Group *AC* (age control) were 94 days old, to provide a control for the difference between Groups *C* and 400 in age and in time on the deprivation-schedule. Ten *Ss* in each of the four groups were tested in one of three conditions: reversal (R), extinction (E), or acquisition (A). In each group, 5 *Ss* were trained with $B+$ and 5 with $D+$ in *OL*. The *Ss* were randomly assigned to groups in blocks of 24, each block completing one replication of the experiment.

(4) Experimental training and testing. Each S was run approximately 23 hr. after its last feeding. Twenty trials per day were given in *OL*, with an intertrial interval of 24 sec. The spatial position of $S+$ was alternated randomly; reward was a 0.25-gm. wet-mash pellet. A noncorrection method was used. The *Ss* were run to a criterion of 18/20 correct on one day. Test conditions (reversal, extinction, or acquisition) were introduced on the day after *OL* was completed. In *RL*, training continued until S reached a criterion of 18/20 correct in one day. In extinction and in acquisition, there were 20 forced trials on each of six days with only one alley open (the other alley was blocked by a translucent door). For extinction-groups, the open alley was illuminated with the *previously rewarded* brightness value (PS^+) and *all trials were nonrewarded*. For acquisition-groups, the open alley was illuminated with the *previously nonrewarded* brightness value, (PS^-) and *all trials were rewarded*. The spatial positions of D and B were alternated randomly.

Results: (1) Original learning. A preference for D was shown on the first trial of *OL* (0.67 of the *Ss* chose dark, $z = 3.82$, $p < 0.0001$). The mean number of days to criterion were 5.0 for the $D+$ problem and 10.4 for the $B+$ problem ($t = 9.92$, $df = 118$, $p < 0.0001$). There were no significant differences between the groups trained on B+ or between those trained on D+.

(2) Reversal learning (RL). Only the results for days to criterion in *RL* will be reported (the results for errors *RL* were essentially the same). Since there was a difference between $D+$ and $B+$ groups in *OL*, the groups were separated for analyses of covariance of days *RL*. The concomitant variable was errors *OL*. In the covariance analyses, the difference between slopes within groups was not significant ($Fs < 1.00$). The mean

number of days to criterion in *RL* (adjusted by covariance) for groups *C-R*, *AC-R*, 200-*R*, and 400-*R* were 13.1, 12.0, 14.2, and 8.3, respectively, for groups with $B+$ in *OL*, and 14.0, 15.6, 19.0, and 14.4, respectively, for groups with $D+$ in *OL*. The difference between groups with $B+$ in *OL* is significant ($F = 7.89$, $df = 3/15$, $p < 0.005$). There is a significant difference between 400-*R* and *AC-R* ($F = 8.25$, $df = 1/15$, $p < 0.025$). It can be concluded that there is a facilitative effect of overtraining *per se; i.e.* the effect is not attributable to differences in age. For groups with $D+$ in *OL*, there is no significant difference in mean number of days *RL* ($F < 1.00$). In contrast to the $B+$ groups, there is no *ORE* with $D+$ in *OL*.

(3) Extinction of response to PS+. To permit later comparisons, the extinction groups were also separated on the basis of $B+$ or $D+$ in *OL*. The results for starting speed in extinction (and in acquisition) will not be reported; the curves showed the same trends as the curves for running speed ($RS = 10 \times$ reciprocal running time). The groups, extinction and acquisition, were combined for analyses of variance of *RS* on the last day of *OL*. There were no significant differences between groups (*C*, *AC*, 200, and 400) in *RS* at the end of training on $B+$ ($F = 1.86$, $df = 3/36$, $p > 0.10$). At the end of training on $D+$ however, there were significant differences in *RS* ($F = 4.10$, $df = 3/36$, $p < 0.025$). Apparently, the ease of the $D+$ problem permits *S*s to reach criterion in *OL* (in terms of choice behavior) before *RS* has reached asymptote.

Initially, the mean *RS* for the last day of *OL* and each day of extinction were plotted for the $D+$ and $B+$ groups. The extinction-curves had two phases: there was a decrease in speed for the first few days, followed by an increase over the last few days. The increase was significant in the $D+$ groups, possibly reflecting the reinforcing properties of the dark alley (which was preferred in *OL*). A decision was made to compare the extinction-curves over the first three days. The curves for groups with $B+$ in *OL* are shown in Fig. 1. There is a significant decrease in *RS* over days ($F = 77.15$, $df = 3/48$). The Groups X Days interaction is significant ($F = 2.14$, $df = 9/48$, $p < 0.05$), indicating that the extinction-curves are not of the same form. It appears that 400-*E* extinguishes faster than the other groups, and reaches a lower level. Although 400-*E* has the lowest mean *RS* over all three days of extinction, the difference between groups is not significant ($F < 1.00$). The difference between groups in linear trend does not reach significance ($F = 2.50$, $df = 3/16$, $p < 0.10$). The extinction-curves for the corresponding groups with $D+$ in *OL* showed essentially the same trends as the $B+$ groups. There was a significant Groups X Days interaction ($F = 3.46$, $df = 9/48$, $p < 0.01$), and the difference between groups in linear trend was significant ($F = 7.43$, $df = 3/16$, $p < 0.005$).

In both the *B*+ and *D*+ groups, 200-*E* and 400-*E* start extinction at higher speeds than *C-E* and *AC-E*. The rank order correlation (ranging from +0.50 to +0.90) between initial *RS* and rate of extinction *within* each of the groups indicates that the *Ss* with high initial speeds tend to extinguish faster than those with low initial speeds. Since initial *RS* is correlated with rate of extinction, and overtraining changes *RS*, the effect of overtraining on the rate of extinction may be due to its effect on initial *RS*.

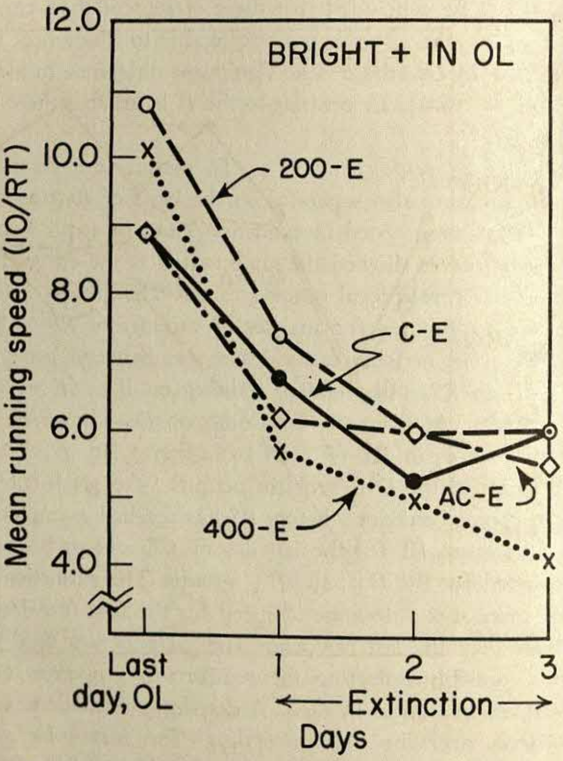


FIG. 1. MEAN RUNNING SPEED ON THE LAST DAY OF *OL* AND THE FIRST THREE DAYS OF EXTINCTION FOR THE GROUPS TRAINED WITH BRIGHT + IN *OL*

(4) Acquisition of response to PS⁻. The curves of acquisition for groups with *B*+ in *OL* are shown in Fig. 2. There is a significant decrease in *RS* from the last day of *OL* to the first day of acquisition ($F = 66.56$, $df =$

1/16). There is a larger decrease in *RS* for the overtrained groups, but the Groups X Days interaction does not reach significance ($F = 3.01$, $df = 3/16$, $p < 0.10$). From the first to the sixth day of acquisition there is a significant increase in *RS* ($F = 34.46$, $df = 5/80$) and a significant Groups X Days interaction ($F = 1.98$, $df = 15/80$, $p < 0.05$). The criterion-groups reach asymptote earlier than the overtrained groups, and the continued increase in speed in the overtrained groups from Day 3 to Day 6 leads to the significant interaction of Groups X Days. For groups with *D+* in *OL*, there was no indication of a larger decrement in *RS* for the overtrained than for the criterion groups from the last day of *OL* to the first day of acquisition. The curves for the four groups were essentially parallel throughout the course of acquisition (*Fs* < 1.00 for the Groups X Days interactions). There is no evidence in the acquisition

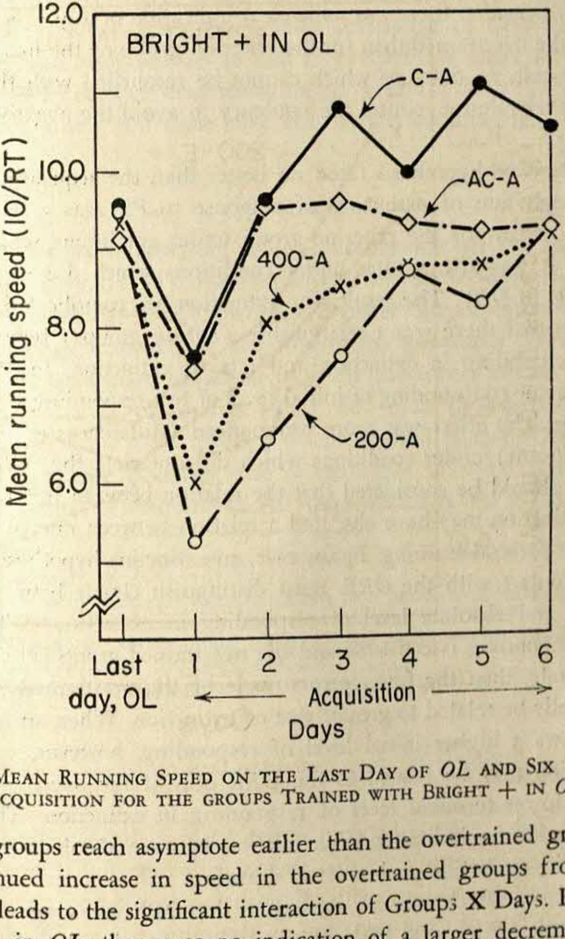


FIG. 2. MEAN RUNNING SPEED ON THE LAST DAY OF *OL* AND SIX DAYS OF ACQUISITION FOR THE GROUPS TRAINED WITH BRIGHT + IN *OL*

curves that the overtrained groups had less avoidance of the original $S-$ than did the criterion groups.

Discussion. The results of this experiment provide no support for the acquisition interpretation of the *ORE*. Under conditions which yielded the *ORE* (400 trials of overturning with $B+$ in OL), there was no suggestion of greater rate of acquisition of response to PS^- in the overtrained group. Furthermore, there was some clear indication of a greater decrement in *RS* for the overtrained than for the criterion-group on the first day of the acquisition-test, an outcome which cannot be reconciled with the assumption that overtraining reduces the tendency to avoid the negative stimulus in training.

The extinction-hypothesis fared no better than the acquisition-hypothesis. Although rate of extinction of response to PS^+ was greater for the overtrained than for the criterion-group under conditions which yielded the *ORE*, it was also greater under conditions which did not yield the *ORE* ($D+$ in OL). The results for extinction are complicated somewhat by the fact that there was a relationship (*within* groups) between initial level of responding in extinction and rate of extinction. In effect, then, there was some confounding of initial level of test-responding with amount of training. The effect was more pronounced (differences in initial level were significant) under conditions which did not yield the *ORE*, and the possibility should be considered that the relation between initial level and rate of extinction may have obscured a relation between rate of extinction and rate of reversal-learning. In any case, an extinction-hypothesis, in order to make contact with the *ORE*, must distinguish clearly between rate of extinction and absolute level of responding in extinction. When initial level of responding is the same and the overtrained group extinguishes at a greater rate, then the fewer errors made by the overtrained group may meaningfully be related to greater rate of extinction. When an overtrained group shows a higher initial level of responding, however, greater rate of extinction may be related to the *ORE* only if the overtrained group reaches a lower terminal level of responding in extinction. The present results provide no indication of a relation between absolute level of responding in extinction and reversal learning.

Two factors which may be related to the *ORE* are suggested by the present results. First, 400 trials of overtraining produced faster *RL* only after training on $B+$, which was the more difficult problem. Although measures of difficulty are not directly comparable across studies, prior studies indicate that problem difficulty may be related to the magnitude of the *ORE*. The largest *ORE* (in terms of percent-savings) was found by Reid,[6] and the original problem was relatively difficult (indicated by trials to criterion in *OL*). Erlebacher,[7] using a relatively easy problem, found no *ORE*. Mackintosh has found some evidence that the magnitude of the overtraining effect will vary with the complexity of the problem.[8] A second possibility is suggested by the present finding that the more difficult problem also led to greater negative transfer in *RL*. Perhaps overtraining has a larger effect when there is more room for improvement, *i.e.* when the criterion group has a large percentage of negative transfer. There have been many failures to find an overtraining effect with position RL.[9] Positional habits of discrimination are generally easier to learn than habits of brightness and there may also be less negative transfer in positional *RL*.

SUMMARY

Rats were trained to criterion, or given 200 or 400 trials of overtraining on a bright (*B*)-dark (*D*) discrimination. Half the *S*s were trained with $B+$ and half with $D+$ and then tested under one of three conditions: reversal-learning, extinction of response to the previously positive stimulus (PS^+), or acquisition of response to the previously negative stimulus (PS^-). With $B+$ in original learning, 400 trials of overtraining gave clear evidence of the *ORE*. With $D+$ in original learning, no evidence of the *ORE* appeared. Neither rate of extinction of response to PS^+ nor rate of acquisition of response to PS^- were related to the *ORE*.

---

[6] L. S. Reid, The development of noncontinuity behavior through continuity learning, *J. exp. Psychol.*, 46, 1953, 107–112.

[7] Albert Erlebacher, Reversal learning in rats as a function of percentage of reinforcement and degree of learning, *J. exp. Psychol.*, 66, 1963, 84–90.

[8] N. J. Mackintosh, Overtraining, reversal, and extinction in rats and chicks, *J. comp. physiol. Psychol.*, 59, 1965, 31–36.

[9] S. E. Sperling, Reversal learning and resistance to extinction, *Psychol. Bull.*, 63, 1965, 281–297.

LEFT-RIGHT FIELD DIFFERENCES WITH PARTIAL REPORT OF LETTERS

By ROBERT E. FITZGERALD and A. J. MARSHALL,
University of Western Australia

When horizontally displayed patterns of letters or symbols are presented in a tachistoscope on one or both sides of a central fixation-point, differences in accuracy of report of the stimulus-elements occur between left and right fields. The side which shows superiority depends on whether the pattern is presented symmetrically (*i.e.* on both sides of fixation simultaneously) or asymmetrically (*i.e.* successively on either side of fixation). Symmetrical presentation generally results in greater accuracy on the left, while asymmetrical presentation results in superior accuracy on the right.[1]

The conditions of symmetrical and asymmetrical presentation involve both different retinal regions and different requirements for report. The letters to be reported come from both sides of fixation in the symmetrical condition only. It follows that the different results obtained under the two conditions might be explained in terms of either the positioning of the pattern or the type of report.

Attempts to account for left-right field-differences have tended to ignore possible inter-relations between the symmetrical and asymmetrical conditions, and have concentrated on independent explanations of the two.[2] Heron claimed that attentional processes could account for both effects.[3] The *S* begins with central fixation but after the exposure there is a tendency to attend to the presented material, reading from left to right. This movement follows more easily when the stimulus items are all to the right of fixation rather than all to the left, and since within each stimulus pattern the left items are reported first these tend to be reported more correctly. Winnick and Dornbush gave verbal instructions for whole or partial (left or right) reporting of letters and found evidence of the influence of attentional factors affecting left-right field differences obtained between pre- and post-exposure instructions.[4]

---

* Received for publication October 18, 1966.

[1] Woodburn Heron, Perception as a function of retinal locus and attention, this JOURNAL, 70, 1957, 38–48; M. P. Bryden, Tachistoscopic recognition of non-alphabetical material, *Canad. J. Psychol.*, 15, 1961, 220–225; Mortimer Mishkin and D. G. Forgays, Word recognition as a function of retinal locus, *J. exp. Psychol.*, 43, 1952, 43–48; H. S. Terrace, The effect of retinal locus and attention on the perception of words, *J. exp. Psychol.*, 58, 1959, 382–385.

[2] M. P. Bryden and C. A. Rainey, Left-right differences in tachistoscopic recognition, *J. exp. Psychol.*, 66, 1963, 568–571; D. S. Camp and E. R. Harcum, Visual pattern perception with varied fixation locus and response recording, *Percept. mot. Skills*, 18, 1964, 287–296.

[3] Heron, *op. cit.*, 38–48.

[4] W. A. Winnick and R. L. Dornbush, Pre- and post-exposure processes in tachistoscopic identification, *Percept. mot. Skills*, 20, 1965, 107–113.

The purpose of the present experiment is to determine whether the left superiority normally obtained with symmetrical presentation and the right superiority with asymmetrical are due to differences in retinal regions or in report requirements. It is proposed to use a technique developed by Sperling of giving a tone signal to obtain independent reports of the left and right halves of a symmetrically presented pattern.[5] Under this condition, the retinal positioning corresponds to that in symmetrical presentation, and the form of report corresponds to that in the asymmetrical condition. It is hypothesized: (1), if retinal positioning is the dominant variable, that accuracy will be greater on the left; and (2), if form of report is dominant, that accuracy will be greater on the right.

Method. Three conditions were employed. In two conditions, horizontal patterns of eight letters were presented with four on either side of the center, and *S* was given a signal, after presentation, to report the left or the right side. In one of these the letters were all evenly spaced, while in the other there was a gap between the fourth and fifth letters. The latter condition was included since a preliminary study suggested that partial report was more efficient when the halves to be reported were separated in this way. The third condition required a full report of four evenly spaced letters presented across the center of the field, and subtending the same visual angle as the eight letters in the previous conditions. This constituted a control condition against which to compare the other two.

Apparatus and materials. A Scientific Prototype Model GB three channel tachistoscope was used to present the patterns. The sequence of fields was Blank, Field 1, Field 2, and return to Blank. The Blank field contained a small black fixation-cross on a matt white ground, Field 1 presented the stimulus-patterns for 300 m.sec. and Field 2, the post-exposure field, comprised a matt black card illuminated for 3 sec. The luminances of these fields were approximately 2.5, 2.5 and 8.0 foot lamberts respectively. The distance from the eyes to Field 1 was 49 in., and to the other two, 48 in. Coupled to the Field 1 output gate was an electrical signal generator which produced immediately after Field 1 went off a high or low pitch tone through a loudspeaker. Selection of the pitch of the tone was effected by a two-way toggle switch. An electronic timer limited the duration of the tone to 40 m.sec.

The stimulus-patterns were typed with an IBM Executive electric typewriter on a roll of white paper. There was a gap of 4.0 mm. between all letters in the even-space eight-letter condition, a 2.3 mm. gap with a 14.3 mm. gap between letters four and five in the second eight-letter condition and a 14.3 mm. gap between all letters in the four-letter condition. The patterns appeared horizontally in the exact center of the field, and the total visual angle under each condition was 2.22°. Only consonants which subtended the same visual angle (letters 3.0 mm. sq.) were employed. This eliminated the letters A E I J M O Q S U W. The letters appearing in each pattern were drawn up at random with no restrictions on the number of times each letter could appear in any one pattern.

Subjects. The *S*s were 30 volunteers from an introductory course in psychology.

---

[5] George Sperling, The information available in brief visual presentations, *Psychol. Mongr.*, 74, 1960, (Whole No. 498), 1–29.

They were screened to ensure that they were English-speaking for several years, and that they had at least 13/16 vision on the near acuity tests of the Bausch and Lomb Ortho-rater both binocularly and monocularly without the aid of glasses.

Procedure. Each *S*s was assigned at random to one of the three conditions. After testing visual acuity, the following instructions were read to the *S*s:

> *Instructions.* In this apparatus I am going to show you a group of eight [four] letters for 0.3 sec. Your task will be to get as many of them correct as you can. No vowels will be included and any letter may appear more than once. If you look through this viewer with both eyes you will see a black cross on a white ground. Try to fixate on this before each set of letters is shown. The letters will lie in a straight line with four [two] on each side of the cross, so that fixating on the cross will help you to get more correct. When you are ready, present the patterns to yourself by pressing this button once. [*Partial report groups only*]: Just at the moment the pattern goes off a tone will sound. A high tone means report only the right four letters. A low tone means report only the left four. [*These were demonstrated to the Ss.*] Make your response immediately by writing as many letters as you can remember, in any order on the response-sheet. Use one space on the response-sheet per pattern. The first 10 patterns will be practice. Any questions? Remember to fixate on the center cross, and to make your response immediately as you will forget the letters very quickly.

Ten practice-trials were given every *S*. During these trials the *S*s were told their score and given the correct response. They then proceeded at their own pace for a further 40 trials.

Scoring. Responses were scored for correct identification of letters without regard to position relative to other letters. If a letter appeared twice in a pattern but was recorded only once, it was scored as one correct, except where it had been presented both on the left and right of fixation in the four-letter pattern when it was scored as one correct on each side. In the two partial-report conditions, each *S* received a score for total correct with partial-left and with partial-right reporting, while in the full-report condition, separate totals for the left and right halves of the patterns were calculated.

Results. The means and standard deviations of the number of letters correct on the left and right of fixation and the difference in accuracy (right-left) between fields under each of the treatment conditions are given in Table I.

The significance of these data was tested using three two-factor analyses of variance with repeated measures on one of the factors.[6] The analyses comparing the Full-report with the Partial-report showed that both Conditions × Positions were significant at well beyond the 0.001 level. The plots of these interactions in Fig. 1 clearly indicate that the accuracy of the Full-report is very much greater on the left side than on the right. The reverse is true for the Partial-reports. The Conditions X Positions interaction for the two Partial-report conditions is also highly significant ($p < 0.005$). Fig. 1 shows that the magnitude of the right field superiority is increased when there is a gap between the four letters on the left and right. There

---

[6] B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 302.

TABLE I

MEANS AND STANDARD DEVIATIONS OF NUMBER OF LETTERS CORRECT IN EACH OF THE CONDITIONS

Condition	Position	Means	*SD*	Mean Difference (Right-Left)
Full-report	left	64.5	11.94	−20.8
	right	43.7	15.73	
Partial-report without gap	left	26.0	10.45	+16.2
	right	42.2	14.04	
Partial-report with gap	left	20.6	8.57	+29.6
	right	50.2	6.99	

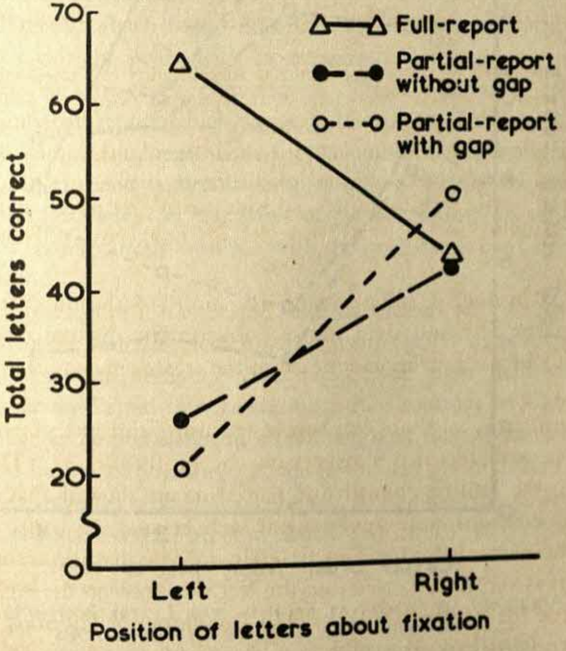


FIG. 1. ACCURACY OF REPORT OF LETTERS FALLING ON THE LEFT AND ON THE RIGHT OF FIXATION AS A FUNCTION OF FULL AND PARTIAL REPORT

is no difference between the total over-all scores in the two Partial-report conditions, but both of these are significantly less than the total score

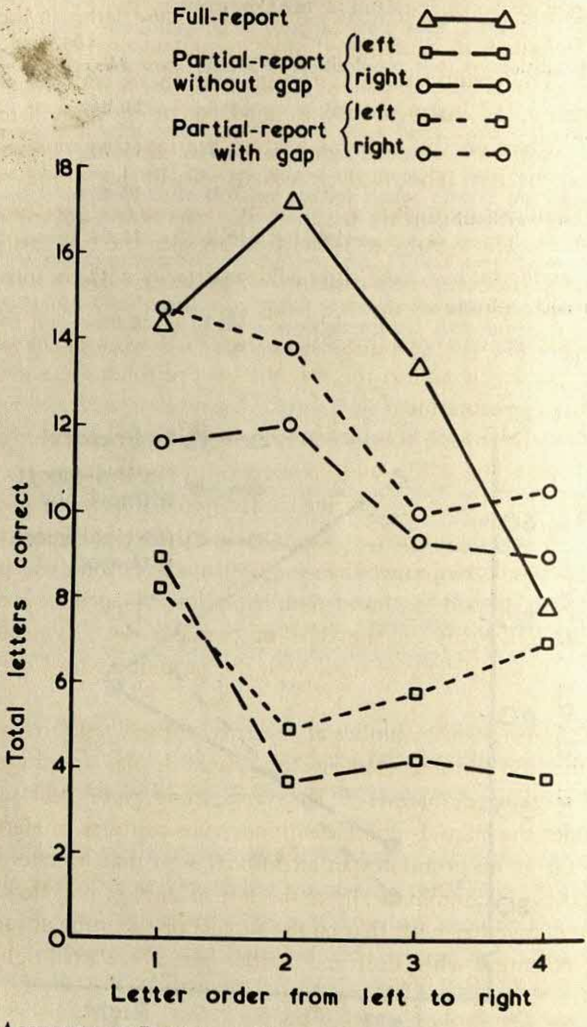


FIG. 2. ACCURACY OF REPORT AT EACH OF THE LETTER POSITIONS (FROM LEFT TO RIGHT) FOR EACH OF THE CONDITIONS

obtained with Full-report ($p < 0.001$). This last result is not relevant to the present study and is almost certainly due to letters in the Full-report condition being much more readily discriminable.

Fig. 2 gives accuracy of report at each of the letter positions for each of the conditions. Under both Partial-report conditions accuracy at all four letter positions on the right is greater than that at any of the letter positions on the left. There is also a tendency for the left-hand letter in the part of the pattern signalled, to be reported more accurately than the other three.

Discussion. The results are in agreement with the expectation that with partial reporting, as instructed by a signal occuring immediately after presentation, letters on the right of fixation are reported more accurately than those on the left. This finding is analagous to that already reported in the literature of right superiority with asymmetrical presentation of groups of letters. The significant interaction between the two partial report conditions showing an increased right field superiority with the introduction of a gap between letters in the two fields supports this conclusion. When the sub-groups are separated by a gap, the two halves are more readily reported independently and in this way the gapped condition is even closer to asymmetrical presentation.

Winnick and Dornbush found no difference between accuracy of report in the left and right fields with post-exposure instructions for partial-report.[7] They also, however, obtained no difference between total scores for partial and full-report which suggests that the verbal instructions for report were delayed for longer than the 250 m.sec. beyond which the advantage of partial report is almost completely lost.[8] Hence the comparison of right with left fields on the basis of partial report, as made in the present study, is not strictly possible with the procedures used by Winnick and Dornbush.[9]

Empirical findings of a number of research workers tend to support the hypothesis that the stimulus-elements in a visual display are processed in a sequentially organized manner.[10] The very strong right field superiority obtained under the Partial-report conditions runs contrary to Heron's proposal in so far as he postulates an attentional scan that operates after the exposure and begins automatically at the left of any set of letters.[11] These data, however, together with that on the success of reporting at each of the four letter positions with Full and Partial-left and Partial-right report, support Bryden and Rainey's conclusion that left field superiority with symmetrical presentation is due to the fading of the memory trace as the pattern is reported in a left to right direction.[12] This implies that a major variable giving rise to the results reported in the literature of left superiority with symmetrical presentation and right superiority with asymmetrical, is that the letters to be reported come from both sides of fixation in the symmetrical condition only.

---

[7] Winnick and Dornbush, *op. cit.*, 107–113.

[8] Sperling, *op. cit.*, 1–29.

[9] Winnick and Dornbush, *op. cit.*, 107–113.

[10] Mishkin and Forgays, *op. cit.*, 43–48; E. R. Harcum, R. R. Hartman, and N. F. Smith, Pre- versus post-knowledge of required reproduction sequence for tachistoscopic patterns, *Canad. J. Psychol.*, 17, 1963, 264–274; Harcum and Smith, Effect of pre-known stimulus-reversals on apparent cerebral dominance in word recognition, *Percept mot. Skills*, 17, 1963, 799–810.

[11] Heron, *op. cit.*, 38–48.

The difference here obtained between left and right partial reports could be explained in terms of either (a) an initial sequential scanning of the letters from left to right giving advantage to those on the right at the time of the post-exposure signal, or (b) time loss on the left since in this case attention must shift from the center to the left before sequential scanning may begin. Either account would mean that in addition to the reporting being increasingly more difficult as the initial impression fades, there is the further factor of the distance from fixation of the first letter that is to be reported.

SUMMARY

Eight letters were exposed horizontally across fixation for 300 m.sec. in a tachistoscope. A signal tone sounding simultaneously with the offset of the exposure-flash instructed *Ss* to report the letters on the left or the right of fixation. Letters on the right were reported more accurately than those on the left. A control condition with four letters, but with all to be reported, showed a strong left field superiority. This finding of right field superiority with partial report is analagous to that obtained when letters are presented in one field at a time. It was concluded that a major factor determining the field differences in the perception of letter patterns reported in earlier studies is whether the letters to be reported come from one or both sides of fixation, irrespective of the positioning of the letters in the stimulus-pattern.

---

[12] Bryden and Rainey, *op. cit.*, 568–571.

DRIVE AND REWARD IN AVERSIVE LEARNING

By WALLACE R. MCALLISTER and DOROTHY E. MCALLISTER,
Syracuse University

Relatively little systematic work has been done on the combination of drive and reward in aversive, instrumental learning, although the appetitive case has been dealt with rather extensively.[1] This state of affairs may be the result of the common tendency to use the complete termination of noxious stimulation as a reward in aversive learning tasks. As a consequence, since drive is normally defined in terms of the strength of the noxious stimulation and reward in terms of the amount of its offset, the independent manipulation of the two variables is prevented. Another likely contributing factor is the absence of a coherent theoretical framework for the aversive case. There have been, however, two studies which manipulated both aversive drive and reward independently.[2] Campbell and Kraeling varied drive by changing the strength of shock in a runway and reward by changing the strength of shock in the goal box. Woods, Davidson, and Peters used similar procedures but varied the temperature of water in an alley and goal box. Neither study, however, was oriented toward studying the combination of the theoretical variables underlying drive and reward which is the major interest of the present paper.

To avoid the theoretical complication of having two sources of drive (pain and fear) and two sources of reward (their offsets) which is inherent in the use of shock as the source of drive in a learning situation, a two-stage, acquired-fear paradigm was adopted in the present studies. In the first stage, a series of inescapable shocks was administered in the start box of a hurdle-jumping apparatus. Subsequently, the *S*s were allowed to jump the hurdle to a safe box, no shocks being presented in this second stage regardless of *S*'s behavior. Typically, the motivation for the learning of the hurdle-jumping response is assumed to be fear, classically

---

* Received for publication November 23, 1966. These investigations were supported in part by research grants, MH-02064 and MH-12978, from the National Institute of Mental Health, Public Health Service. The authors are indebted to Joseph J. Franchina, Richard B. Tenser, and W. Keith Douglass who aided in the collection and analysis of the data.

[1] R. W. Black, On the combination of drive and incentive motivation, *Psychol. Rev.*, 72, 1965, 310–317.

[2] B. A. Campbell and Doris Kraeling, Response strength as a function of drive level and amount of drive reduction, *J. exp. Psychol.*, 45, 1953, 97–101; P. J. Woods, E. H. Davidson, and R. J. Peters, Jr., Instrumental escape conditioning in a water tank: Effects of variations in drive stimulus intensity and reinforcement magnitude, *J. comp. physiol. Psychol.*, 57, 1964, 466–470.

conditioned to the stimuli in the start box, and the reward, the decrease in fear occurring following the hurdle-jump.

The above procedure was used in two experiments which differed in drive-level through the use of different intensities of shock in fear conditioning. Drive was also varied within each experiment by the manipulation of the number of fear-eliciting stimuli. High drive depended on fear conditioned to both a discrete *CS* and to the static apparatus cues, while low drive involved fear conditioned only to the latter stimuli. Evidence that these procedures lead to different strengths of drive has been presented previously.[3] At each drive level in each experiment, the magnitude of reward was varied by manipulating, through the use of a novel or familiar safe box, the amount of fear reduction occurring following a response. It was reasoned that escape to a novel safe box would result in less fear reduction than escape to a familiar safe box. This assumption was based on previous reports that novel stimuli can elicit fear[4] and on a recent study by McAllister, McAllister, and Zellner.[5] In this latter study, using the apparatus of the present experiments, the *Ss*, given familiarization in a grid box showed, an increased preference, in a later free-choice situation, for the grid box over the novel safe box as compared to a group not given such training. An analogous strategy for manipulating the amount of reward independent of drive was employed in avoidance conditioning studies by Bower, Starr, and Lazarovitz and by Knapp.[6] Their method of varying the amount of fear reduction involved changing the similarity of the stimuli in the start and safe boxes. These studies, however, did not vary drive level. With the present design, drive-reward combination can be studied at both high and low levels of the variables.

Method. The two studies to be reported here used the same design and procedure except for the level of shock employed during fear conditioning.

Subjects. Naïve, female, hooded rats, ranging in age from 91–122 days, from the colony maintained in the Experimental Psychology Laboratory at Syracuse University, were used. In Experiment I, there were 64 *Ss*, 16 in each of four groups; in Experiment II, 48 *Ss*, 12 in each of four groups. Two additional *Ss* in the second experiment were discarded, one because of illness and one because of difficulty in handling during the training trials. Within each experiment, the *Ss* were randomly paired and then the pair was assigned at random to an experimental condition.

Apparatus. For the fear-conditioning phase of each experiment, two identical white grid boxes, wired in parallel, were used so that *Ss* could be treated in pairs. Hinged to the top of each grid box was another box which contained incandescent lamps serving as light sources for the intertrial illumination (7 ft.-c.) and the discrete *CS* (115 ft.-c.). The *UCS* was a shock which energized the grids successively at a rate of two impulses per grid per sec. A 100,000-Ω resistor was in series with each *S*. For Experiment I, the shock was 125 v.; for Experiment II, 70 v.

---

[3] W. R. McAllister and D. E. McAllister. Variables influencing the conditioning and the measurement of acquired fear, in W. F. Prokasy (ed.), *Classical Conditioning: A Symposium*, 1965, 172–191.

[4] See, for example, D. E. Berlyne, Curiosity and exploration, *Science*, 153, 1966, 25–33; K. C. Montgomery, The relation between fear induced by novel stimulation and exploratory behavior, *J. comp. physiol. Psychol.*, 48, 1955, 254–260.

[5] D. E. McAllister, W. R. McAllister, and D. K. Zellner, Preference for familiar stimuli in the rat, *Psychol. Rep.*, 19, 1966, 868–870.

[6] Gorden Bower, Ronald Starr, and Leah Lazarovitz, Amount of response-produced change in the CS and avoidance learning, *J. comp. physiol. Psychol.*, 59, 1965, 13–17; R. K. Knapp, Acquisition and extinction of avoidance with similar and different shock and escape situations, *op. cit.*, 60, 1965, 272–273.

The apparatus for the hurdle-jumping phase of the experiment consisted of a start box, which was a replica of the grid boxes described above, separated by a guillotine door from a gray safe box with a plywood floor hinged so as to serve as a floor switch. The apparatus, as well as the general procedure, has been described more fully in a previous report.[7]

Design and procedure. Within each experiment, on both of the first two days, two of the groups were handled and were allowed to explore each side of the apparatus. Handling consisted of alternately holding and petting *S* and placing it on a table to explore. The sequence of events for one *S* of a pair was handling, exploration of the grid box of the hurdle-jumping apparatus, handling, and exploration of the safe box for 10 min. each. For the other *S*, the sequence began with exploration of the grid box. The other two groups received the same handling treatment but were returned to their home cage instead of exploring the hurdle-jumping apparatus.

On the third day, 35 fear-conditioning trials were administered with a 2-min. intertrial interval. The conditioning trials for one group given handling and exploration and one group given handling alone consisted of the presentation of a 6-sec. *CS* (increase in illumination from 7 to 115 ft.-c.), the last 2 sec. overlapping the *UCS*. Both stimuli terminated simultaneously. For the other two groups, the onset of the *CS* followed the offset of the *UCS* by 15 sec. These two procedures are, respectively, forward- and backward-conditioning paradigms. To equalize the amount of time spent in the conditioning box, the forward-conditioning *S*s were placed in the boxes 605 sec. before the first trial was begun.

On each of the fourth and fifth days, 25 hurdle-jumping trials, with a minimal intertrial interval of 30 sec., were given to all *S*s in an identical manner. These days will be designated Day 1 and Day 2 of hurdle jumping. No shock was administered during this training. On each trial, *S* was placed in the grid box facing the guillotine door which was raised after 10 sec. simultaneously with the presentation of the *CS* and the activation of a Standard Electric timer, calibrated in 0.01 sec. Depression of the floor switch in the safe box stopped the timer and terminated the *CS*. After 10 sec., *S* was removed to a holding box. If a jump did not occur in 60 sec., *S* was removed to the holding box and a 60-sec. latency recorded. Training was terminated on the second hurdle-jumping day if *S* failed to jump in 60 sec. on 10 consecutive trials. In this case, a 60-sec. latency was recorded for the remaining trials.

The variables manipulated as described above define drive and reward as used in the present studies. In summary, within each experiment, drive level was defined in terms of the number of fear-eliciting stimuli. Groups given classical, forward-conditioning trials had fear conditioned to both the discrete *CS* and to the static cues of the apparatus and constituted the high-drive (*HD*) condition. Groups given classical, backward-conditioning trials had fear conditioned only to the static apparatus cues and constituted the low-drive (*LD*) condition. Drive was further manipulated between experiments by using different intensities of shock during fear conditioning. Thus, the *HD* and *LD* groups of Experiment I had higher drive levels than their counterparts in Experiment II. Amount of reward was defined in terms of the amount of fear reduction occurring after a jumping response and was experimentally manipulated through the variation in the amount of fear elicited in the safe box. The high amount of safe-box fear defining the low reward (*LR*) condition was obtained by using a novel, unexplored safe box. The low amount of safe box fear defining the high reward (*HR*) condition was obtained by using a familiar, previously explored safe box. It should be pointed out that although the same procedure for manipulating safe-box fear was used in both experiments, the levels of reward were greater in Experiment I since the stronger drive levels used there permitted greater amounts of fear reduction.

---

[7] McAllister and McAllister, Postconditioning delay and intensity of shock as factors in the measurement of acquired fear, *J. exp. Psychol.*, 64, 1962, 110–116.

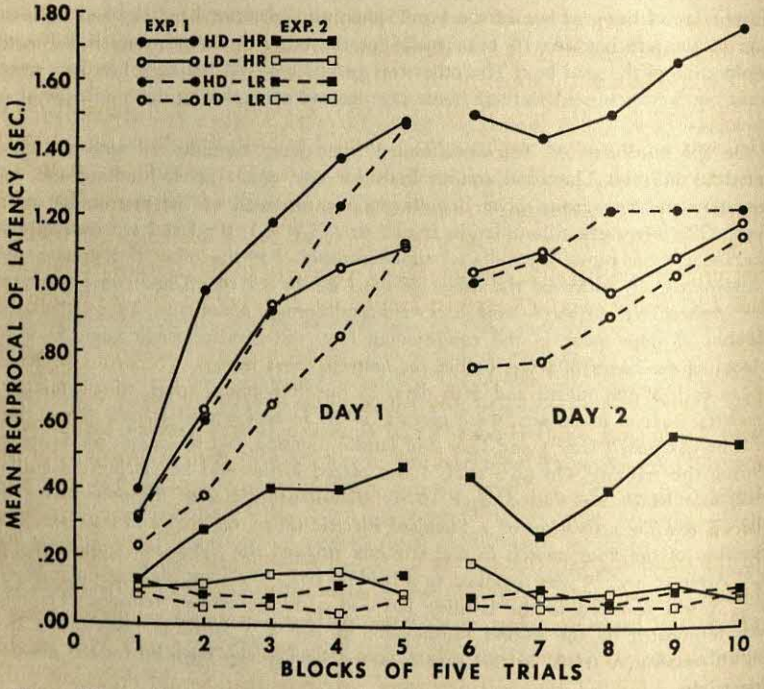


FIG. 1. HURDLE-JUMPING PERFORMANCE FOR HIGH DRIVE (*HD*), LOW DRIVE (*LD*), HIGH REWARD (*HR*), AND LOW REWARD (*LR*) COMBINATIONS IN THE TWO EXPERIMENTS

Results. Mean reciprocals of latency of hurdle jumping in seconds, as a function of blocks of five trials, are presented in Fig. 1 for each of the four groups in each experiment. The upper four curves were obtained in Experiment I while the lower four curves were obtained in Experiment

II. The striking difference in the general levels of performance between the two sets of curves is attributable to the difference in the intensity of shock used during classical fear conditioning in the two experiments. On this basis, each group in Experiment I would have a higher level of fear than the like group in Experiment II. These data corroborate previous findings.[8]

For Experiment I, the data for Day 1 were analyzed with a trend analysis of variance with Drive (D), Reward (R), and Trial Blocks (T) as the factors. Both the $D \times T$ and $R \times T$ interactions were significant ($F = 2.62, 3.50$; $df = 4/240$; $p < 0.05, < 0.01$, respectively). Therefore, factorial analyses of variance were computed for each trial block. None of these analyses revealed a significant interaction. On each trial block, HD was significantly superior to LD ($F = 4.23, 15.00, 6.69, 9.14, 7.70$; $df = 1/60$; $p < 0.05, < 0.001, < 0.025, < 0.005, < 0.01$, respectively). HR was significantly superior to LR only on the first three trial blocks ($F = 5.77, 18.22, 7.69$; $df = 1/60$; $p < 0.025, < 0.001, < 0.01$, respectively). These findings indicate that there was no interaction between D and R. The change in the effectiveness of the R variable over trials is not unexpected because the increasing exposure to the novel safe box for the LR groups would increase familiarity and permit the extinction of fear.

As shown in the figure, a decrement in performance occurred between Day 1 and Day 2 only for the LR groups. A trend analysis of variance with D, R, and Trial Blocks 5 and 6 as the factors yielded a significant $R \times T$ interaction ($F = 10.08$, $df = 1/60$, $p < 0.005$). Analysis of the simple effects showed HR to be significantly superior to LR on Trial Block 6 ($F = 11.85$, $df = 1/60$, $p < 0.005$) but not on Trial Block 5. In the trend analysis there were no significant interactions involving the D factor; performance with HD was significantly superior to that with LD ($F = 11.71$, $df = 1/60$, $p < 0.005$). The pattern of results suggests a spontaneous recovery of safe-box fear between Days 1 and 2. It might, then, be expected that extinction of safe-box fear would occur over the Day 2 trials, as it had on Day 1, with a similar convergence of the curves for the two HD and for the two LD groups. Graphically, this appears to be the case under the LD, but not under the HD, condition. A trend analysis of variance over the Day 2 trial blocks yielded, however, no significant interactions between D, R, and T but each factor was significant ($F = 6.93, 4.33, 6.27$; $df = 1/60, 1/60, 4/240$; $p < 0.025, < 0.05, < 0.001$, respectively).

---

[8] M. L. Goldstein, Acquired drive strength as a joint function of shock intensity and number of acquisition trials, *J. exp. Psychol.*, 60, 1960, 349–358.

Nonparametric statistics were used in the analysis of the data of Experiment II because of the extremely restricted range of scores for the three groups which did not learn the hurdle-jumping response. Mann-Whitney *U* tests (two-tailed), using the mean reciprocal of latency on Trial Block 5, indicated that significantly superior performance ($p < 0.05$) was obtained with *HR* as compared with *LR* under the *HD* ($U = 33$), but not under the *LD* ($U = 67.5$), condition. Also *HD*, as compared to *LD*, led to better performance with *HR* ($U = 25.5$), but not with *LR* ($U = 69$). On Day 2, using the sum of the reciprocals of latency over all trial blocks, the same pattern of results was obtained. Changes in performance from Trial Block 1 to 5 were evaluated with *t*-tests. Only the *HD-HR* group showed a significant improvement ($t = 2.428$, $df = 11$, $p < 0.05$). These findings indicate an interaction between *D* and *R*.

Empirically, the results of the present studies demonstrate that, in this situation, the intensity of shock used for conditioning, the number of fear-eliciting stimuli, and the familiarity of the safe box are relevant variables.

Discussion. Investigations of drive-reward combination have usually used appetitive motivation. A number of such studies, typically involving a factorial design with two levels of drive and two levels of reward, have been reviewed by Black.[9] In general, an interaction between the two variables has been obtained when one level of drive was zero but not when higher levels were employed. These findings appear to be inconsistent with the assumption made by Spence that the theoretical variables underlying drive and reward, *D* and *K* (incentive motivation), combine additively.[10] Black, however, has offered a reconciliation of the empirical findings and the additive hypothesis. He proposes that the strength of the goal response (R_G) and, hence, its conditioned anticipatory form (r_g), the basis of *K*, depends on drive as well as on reward variables. That is, the greater the *D* the greater would be the value of *K*. With zero relevant drive, R_G would not occur and would, thus, yield a zero value of *K* regardless of the presence of a potentially adequate reward. Therefore, either zero relevant drive or zero reward would yield a zero value of *K*. The consequence of these considerations is that, even though *D* and *K* combine additively, an interaction would be predicted in the usual 2 × 2 factorial design when one level of drive or of reward is zero. With higher values of drive and reward, no interaction would be predicted with the additive hypothesis.

---

[9] R. W. Black, On the combination of drive and incentive motivation, *Psychol. Rev.*, 72, 1965, 310–317.

[10] K. W. Spence, *Behavior Theory and Conditioning*, 1956, 196–198.

An analogous theoretical treatment can be tentatively proposed to account for the present results obtained with aversive motivation.[11] Fear reduction or relief occurring after a jumping response may be considered to be the goal response. A conditioned anticipatory form of this goal response, similar in operation to r_g, would presumably develop with trials and serve as a basis for *K*. Of necessity, *K* would be partially determined by drive (fear) since with zero fear, there could be no goal response. It follows that, analogous to the appetitive case, when either fear or relief are zero, the value of *K* would be zero. As applied to the present data, these considerations suggest a basis for the interaction which occurred in Experiment II. If it is assumed that the amount of fear conditioned in each of the *LD* groups was negligible, relief could not occur for either *HR* or *LR* and *K* would be zero. *K* would also be zero under the *HD-LR* condition if the amount of fear elicited in the safe box was equal to or greater than that elicited in the start box since under such a circumstance relief could not occur. On the other hand, relief would be expected to occur with the larger difference in start- and safe-box fear in the *HD-HR* group and result in some value of *K*. The outcome would be a significant interaction, as obtained, even if *D* and *K* combine additively. In Experiment I, where the values of the variables were higher, the results are consistent with the additive hypothesis. It is true that a multiplicative relation between *D* and *K* would also predict the results of Experiment II, but it would not be able to account for the lack of an interaction in Experiment I.

SUMMARY

Following classical fear-conditioning procedures with a 125-v. shock in Experiment I and a 70-v. shock in Experiment II, four groups of rats in each experiment were allowed to jump a hurdle to escape the stimuli paired with shock. Within each experiment, level of drive (number of fear-eliciting stimuli) and magnitude of reward (amount of fear reduction) were varied. Investigation of drive-reward combination revealed an interaction with low levels of the variables (Experiment II) but not with high levels (Experiment I), in agreement with results obtained with appetitive motivation. An incentive-motivation explanation was suggested. The results were interpreted to be consistent with the hypothesis that the theoretical variables, *D* and *K*, combine additively.

---

[11] Other investigators have also considered the operation of a positive incentive in aversive learning situations. See, for example, N. E. Miller, Some reflections on the law of effect produce a new alternative to drive reduction, in M. R. Jones (ed.), *Nebraska Symposium on Motivation*, 1963, 65–112; *passim;* Larry Stein. Facilitation of avoidance behavior by positive brain stimulation, *J. comp. physiol. Psychol.*, 60, 1965, 9–19.

BINAURAL SUMMATION IN LOUDNESS OF TWO TONES AS A FUNCTION OF THEIR BANDWIDTH

By R. D. PORSOLT and R. J. IRWIN, University of Auckland, New Zealand

Binaural summation of loudness refers to the phenomenon that a sound heard with two ears is usually judged louder than the same sound heard with one; or, alternatively, that the sound-pressure of a binaural sound is usually less than that of an equally loud monaural one. This paper advances the hypothesis that binaural summation of loudness can be explained by events taking place in the inner ear, and reports experimental evidence that tends to confirm the hypothesis.

It is hypothesized that binaural summation is related to an increment in loudness that is sometimes observed with monaural stimuli. If a band of white noise, heard monaurally, is increased in width, while its total sound-pressure is held constant, the loudness of the noise remains unchanged until a certain critical width is exceeded; after which its loudness begins to increase with bandwidth.[1] The same effect is observed with combinations of pure tones whose over-all separation in frequency is varied: up to a critical separation, the loudness of the tones is constant for a constant sound-pressure; beyond that separation, loudness increases as a function of bandwidth, even though sound-pressure remains unchanged.[2] Thus monaural sounds whose spectra are wider than a 'critical band' produce a louder effect than those of equal sound-pressure but lesser bandwidth. The width of a critical band is a function of its central frequency,[3] but at any particular frequency it is independent of such factors as sound-pressure,[4] the number of components in the sound,[5] spectral shape,[6] or whether the threshold has been artificially raised by a masking noise.[7]

Monaural summation in loudness as a function of bandwidth may be related to the response of the inner ear. The point of maximal displacement of the basilar membrane has been shown to depend upon the frequency of stimulation;[8] and it

---

* Received for publication November 28, 1966.

[1] Richard Feldtkeller and Eberhard Zwicker, *Das Ohr als Nachrichtenempfänger*, 1956, 81–83; Eberhard Zwicker, Georg Flottorp, and S. S. Stevens. Critical band width in loudness summation. *J. acous. Soc. Amer.*, 29, 1957, 552–554.

[2] Feldtkeller and Zwicker, *op. cit.*, 79–81; Zwicker, Flottorp, and Stevens *op cit.*, 548–551; Bertram Scharf, Loudness of complex sounds as a function of the number of components. *J. acous. Soc. Amer.*, 31, 1959, 783–785.

[3] Zwicker, Flottorp, and Stevens, *op. cit.*, 555–556.

[4] Zwicker, Flottorp, and Stevens, *op. cit.*, 550.

[5] Scharf, *op. cit.*, 783–785.

[6] Scharf, Loudness summation and spectrum shape, *J. acous. Soc. Amer.*, 34, 1962, 228–233

[7] Scharf, Loudness summation under masking, *J. acous. Soc. Amer.*, 33, 1961, 503–511.

[8] Georg von Békésy, Variation of phase along the basilar membrane with sinusoidal vibrations, *J. acous. Soc. Amer.*, 19, 1947, 452–460.

therefore seems that stimulating a wider region of the membrane is more effective than stimulating a subcritical region. This explanation may be as equally applicable to two ears as to one. If, for whatever reason, stimulating separated parts of the same membrane is more effective than stimulating one part, stimulating two basilar membranes might be supposed to be more effective than stimulating one, for similar reasons. For example, if the relative ineffectiveness of two neighboring frequencies in a single ear results from their mutual interference, owing to the spatial proximity of the tones' effects on the basilar membrane, then this spatial proximity, and its consequential interference, would not occur if the tones were presented binaurally, one to each ear.

A test of this hypothesis of the origin of binaural summation can be made with two pure tones that can be presented separately to each ear, or together to one. If the two tones are separated by less than a critical bandwidth, then the hypothesis predicts their binaural loudness will be greater than their monaural loudness. When, however, their bandwidth is so supracritical that no mutual interference is experienced either binaurally or monaurally, then our hypothesis predicts that their binaural loudness will equal their monaural loudness. This prediction is made for the condition when the sum of the binaural sound-pressures equals the monaural sound-pressure. For equal sound-pressures in each ear, this match would still constitute 3 db. binaural summation as it is conventionally measured.

An experiment was conducted to discover whether this was so. The combined loudness of two tones, heard monaurally, was matched to the loudness of the two tones heard binaurally, with one tone in each ear. The difference in decibels between the equally loud monaural and binaural stimuli provided a measure of binaural summation. The binaural summation was determined as a function of the separation in frequency of the two tones (their bandwidth). This function was determined for bandwidths about two different central frequencies, and also at two different sound-pressure-levels (*SPLs*) of the binaural stimulus.

This test of the hypothesis rests on the assumption that although monaural loudness increases with bandwidth, binaural loudness does not. Experimental evidence is largely in support of this assumption. Hughes found that the binaural threshold was unaffected by differences in frequency between the ears.[9] Békésy, on the other hand, found that loudness increased most when a tone added to the other ear was of the same frequency.[10] In his experiment, however, observers were instructed to attend to a fixed tone in one ear, and regard the variable frequency as a disturbance that might affect loudness. In a more recent experiment, Niese presented narrow bands of noise separately to each ear and found that the binaural difference in frequency of the bands did not effect their loudness.[11]

METHOD AND PROCEDURE

Apparatus. A simplified schematic diagram of the apparatus is shown in Fig. 1. The tones were produced by two oscillators, O_1 and O_2 (Leader Type *LAG* 55). Each oscillator was adjusted to the desired frequency with the aid of a Wien Bridge circuit. By means of a switch, SW_1, the tones could be sent either (a) separately to each earphone via *T*-pads, T_{1a} and T_{1b}, or (b) together to one earphone after being mixed in a *T*-pad, T_2, and attenuated by an attenuator, *ATT*, and 'sone potentiometer,' *SP*.[12] The observer (*O*) adjusted the loudness of the monaural stimulus by means of the sone potentiometer. The attenuator was so altered from trial to trial that a particular setting of the potentiometer did not correspond to a particular intensity in the earphone.

---

[9] J. W. Hughes, The monaural threshold: Effect of subliminal contralateral stimulus, *Proc. roy. Soc.*, 124B, 1938, 406–420.

[10] Békésy, *Experiments in Hearing*, 1960, 225–227.

[11] Herbert Niese, Subjektive Messung der Lautstärke von Bandpassrauschen, *Hochfrequenztechn. u. Electroak.*, 68, 1960, 213.

The duration of the tones was controlled by a relay, SW_2, which was operated by a cam driven by a motor, *M*. The motor rotated once every 12 sec. during which time four signals were presented in the following sequence: first, two presentations of about 1.25 sec. separated from each other by about 0.5 sec.; then followed a 3-sec. interval, after which the first sequence was repeated, thus completing the 12-sec. cycle of the motor. During the 0.5-sec. interval between presentations, SW_1 could be so switched manually that the two tones could be sent together to one ear or separately to each ear. Because the tones were switched on and off abruptly, transients were produced at the onset and finish of a presentation. A comparatively long duration of presentation of the tones was chosen in an attempt to overcome the effects of these transients.

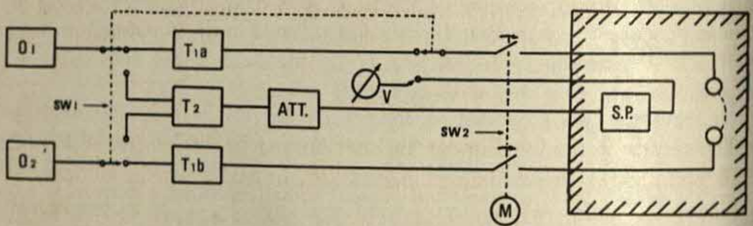


FIG. 1. SCHEMATIC DIAGRAM OF THE APPARATUS

The level of the tones in the earphones was measured with a voltmeter (Brüel and Kjær, Type 2410). Amplivox 82 earphones, used throughout the experiment, were calibrated over the required frequency-range on a 6-c.c. coupler. The *SPL* produced in the earphones by a constant voltage was fairly flat (± 2.5 db.) from 200 to 2500 ~, and the maximal difference between the earphones at a single frequency was 0.5 db.

Additional switches, not illustrated in Fig. 1, allowed O_1 and O_2 to be interchanged, and also allowed the monaural signal to be sent to either ear.

Observers. Sixteen men volunteered to serve as *O*s. All were associated with the Department of Psychology, University of Auckland. Their ages ranged from 20 to 29 yr. Their pure-tone sensitivity was measured with an Amplivox 82 audiometer over the range of frequencies used in the experiment. All *O*s satisfied the criteria of a hearing loss not greater than 10 db. (British Standard) at any frequency in either ear, and a difference in sensitivity between the ears not greater than 10 db. for a particular frequency. One person was not included in the 16, although he satisfied the criteria, because his adjustments proved to be very erratic.

---

[12] E. C. Poulton and S. S. Stevens, On the halving and doubling of the loudness of white noise, *J. acous. Soc. Amer.*, 27, 1955, 329.

Procedure. An *O*'s task was to adjust the loudness of two tones heard monaurally to equal their loudness heard binaurally, with one tone in each ear. The experiment was divided into two parts, with eight different *O*s in each part. The first part required monaural matches to binaural stimuli at an *SPL* of 80 db. (*re* 0.0002 dyne/cm^2) in each ear (a total of 83 db. *SPL*), and the second part required matches to binaural stimuli of 60 db. *SPL* in each ear (a total of 63 db. *SPL*).

Within each part, the *O*s matched pairs of tones separated by seven different bandwidths, and centered about two different frequencies. One central frequency was 500 ~ it was the arithmetic mean of tones whose bandwidths were 0, 40, 80, 120, 160, and 400 ~. The other central frequency was 2000 ~ it was the arithmetic mean of tones whose bandwidths were 0, 100, 200, 300, 400, 500, and 1000 ~ These bandwidths were chosen in the expectation that they would cover a range that was both subcritical and supercritical at central frequencies of 500 and 2000 ~ respectively.

Before the experiment began, each *O* was given practice at making judgments of equality until he seemed to be judging stably. In the experiment, each *O* made eight judgments for every combination of bandwidth and central frequency. Two 2-hr. sessions were required of each *O*; during a session, an *O* made 56 matches for bandwidths round one central frequency.

RESULTS

The results of the experiment are summarized in the graphs of Fig. 2. Each point shows the median monaural *SPL* judged equal in loudness to a binaural stimulus whose tones were separated by a given bandwidth (ΔF), round a given central frequency (C.F.), at a given *SPL*. The vertical lines show the interquartile ranges of the 64 judgments on which each median is based. The dashed horizontal lines show the total binaural level, which is 3-db. greater than the level in either ear. All points above this line, therefore, indicate that the monaural stimulus was adjusted to a level greater than the sum of the binaural levels in order to make it sound equally loud.

The reliability of the effects depicted in Fig. 2 was evaluated by analysis of variance. Each *O*'s average summation (in db.) for eight judgments at a given combination of bandwidth, central frequency, and binaural *SPL*, was the basic data for the analysis.

The analysis showed that the decline in binaural summation as a function of bandwidth that is exhibited in Fig. 2, and predicted by our hypothesis, was significant ($P < 0.001$). The decline exhibited for bandwidths round 500 ~ conform qualitatively with the hypothesis, in that narrow bands give generous summation over the total binaural *SPL*, and wide bands give none at all. But the decline with bandwidth exhibited for a central frequency of 2000 v poses a problem, because the summation falls below the horizontal lines. This difference in over-all summation at the two central frequencies is significant ($P < 0.001$). One possible explanation, though one which our experiment cannot assess, rests on the fact that the greatest binaural summation is obtained when the sensation-levels in each

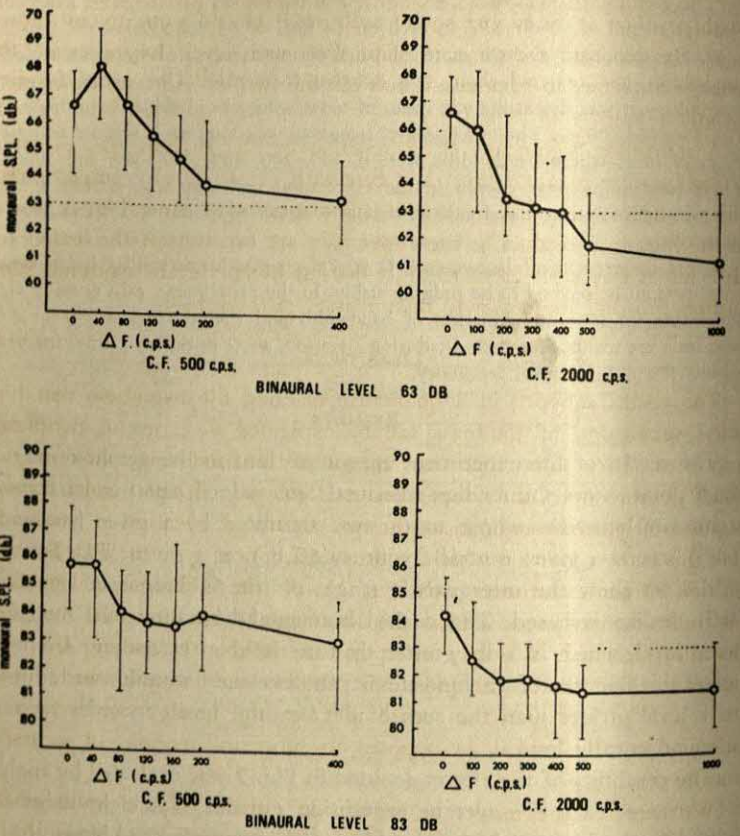


FIG. 2 BINAURAL SUMMATION OF TWO TONES AS A FUNCTION OF BANDWIDTH (ΔF)

Each point is the median of 64 monaural-binaural matches (8 by each of 8 Os), and the vertical lines are interquartile ranges. The dotted horizontal lines show the total binaural *SPL* (3 db. greater than the level in each ear).

ear are the same.[13] The wide bandwidths employed round 2000 ~ might have given opportunities for differences in sensation-levels between the ears to become appreciable at different frequencies, and thus diminish the amount of binaural summation obtained.

---

[13] R. J. Irwin, Binaural summation of thermal noises of equal and unequal power in each ear, this JOURNAL, 78, 1965, 57–65.

More binaural summation was found with a total binaural *SPL* of 63 db. than with binaural total of 83 db. ($P < 0.05$). This result is inconsistent with a common finding that binaural summation increases with level.[14] There are suggestions in some published data that the increase might level off at the higher sound-pressures,[15] but no work, to our knowledge, has reported a decline in summation with level. In any event, the average difference in summation between the two levels in our experiment is only 1.5 db.

A remaining significant effect revealed by the analysis was an interaction between binaural level and bandwith ($P < 0.001$). Inspection of Fig. 2 suggests that the decline in summation with bandwidth is slower at the higher level than at the lower. We are not sure of the reason for this, but it may stem from the fact that, at higher levels, monaural summation proceeds more slowly than at moderate levels.[16]

DISCUSSION

The results support, in a qualitative manner, the hypothesis that binaural summation of loudness can be explained in terms of peripheral processes. They show that the amount of binaural summation of two tones depends upon their separation in frequency. In the normal demonstration of binaural summation, the two stimuli are identical in frequency, and this is the condition which, with one exception, gave the greatest summation in our experiment (see Fig. 2). As the difference in frequency of the two tones increased, so the advantage of binaural over monaural listening declined. We may interpret this decline as resulting from the increased loudness of the monaural stimulus—an increase which, in its turn, can be interpreted as resulting from the greater effectiveness of stimulating supracritical as opposed to subcritical regions of a basilar membrane.

We expected the results to provide an estimate of the width of the critical bands at 500 and 2000 ∼. In fact, however, Fig. 2 reveals no discontinuities that can be used to provide such estimates. If one were prepared to assume that the critical band has sharply defined limits, it would be possible to fit to the points of Fig. 2 pairs of interesting straight lines (one horizontal, the other declining from left to right) that might seem to summarize the effect of bandwidth on binaural summation. But this arbitrary assumption about the shape of the critical band would yield correspondingly arbitrary estimates of its width. Consequently no estimates have been made.

---

[14] G. S. Reynolds and S. S. Stevens, Binaural summation of loudness, *J. acous. Soc. Amer.*, 32, 1960, 1337–1344.

[15] Harvey Fletcher and W. A. Munson, Loudness, its definition, measurement and calculation, *J. acous. Soc. Amer.*, 5, 1933, 82–108; B. Caussé and Paul Chavasse, Différence entre l'écoute binauriculaire et monauriculaire pour la perception des intensités supraliminaires, *C. R. Soc. Biol.*, Paris, 136, 1942, 405–406; Irwin, *op. cit.*, 60–63.

[16] Zwicker, Flottorp, and Stevens, *op. cit.*, 551.

The hypothesis that binaural summation is similar in origin to monaural summation of loudness requires qualification. Monaural summation is typically greatest at moderate levels, *i.e.*, 50-60 db. *SPL*, whereas binaural summation usually increases with level. While this difference means that monaural and binaural summation will not necessarily be quantitively identical, it does not destroy the possibility that they may both have similar peripheral origins; or, to put it less paradoxically, that the difference between monaural and binaural loudness can be explained in terms of a lack of monaural summation which is a peripheral phenomenon. The relation between monaural summation and stimulus-level has been explained in terms of the hypothetical relation between the spread of stimuli over the basilar membrane and stimulus-level, and thus in terms of the mutual interference of stimuli.[17] These considerations do not apply to the binaural condition, and the same relation between summation and stimulus-level would not necessarily be expected. For example, in an unpublished study of the summation of two bands of noise, adjacent in frequency, we failed to obtain as much monaural as binaural summation. It was doubtless unrealistic to expect identical summation in these circumstances, for identical summation could only be expected for stimuli whose level and separation in frequency generated maximal monaural summation.

SUMMARY

The hypothesis was advanced that binaural summation of loudness is similar in origin to the increment in monaural loudness that results when the bandwidth of a stimulus of constant sound-pressure is increased beyond a critical value. Sixteen *O*s therefore adjusted the level of two tones heard monaurally to equal their loudness when one tone was presented to each ear. The adjustments were made with the two tones at various bandwidths. The results were qualitively consistent with the hypothesis: as the bandwidth of the tones increased, binaural summation declined. Under some, but not all, conditions equal binaural and monaural loudnesses corresponded to equal binaural and monaural sound-pressures; the limited circumstances under which this equality would be expected to hold were discussed.

---

[17] Eberhard Zwicker and Bertram Scharf, A model of loudness summation, *Psychol. Rev.*, 72, 1965, 12–16.

RULE LEARNING IN DEAF AND HEARING SUBJECTS

By PENELOPE B. ODOM and RICHARD L. BLANTON,
Vanderbilt University

The hearing child acquires most of his knowledge of language through audition, a sense which processes sequential information. Since he has acquired most of his language before he can read or write, he has only temporal positional and relational cues as to the usage and often the semantic content of a new element heard for the first time. The deaf child, however, does not have access to this source of input and must, therefore, learn language by its visual characteristics. The most commonly used language of the deaf, Sign, is quite unlike spoken English in that word-order is not as rigid, neither is it a highly inflected language.[1] A language without either of these characteristics would be likely to provide the learner with less information about its components and unfamiliar elements than a language which was inflected or had rigid word order requirements. That is, there would be few cues to reference and part of speech from position in the utterance, only those inherent in the sign itself. It might be expected, then, that the deaf might have difficulty remembering the temporal order of unfamiliar material, because it is not a dimension that has been relevant in their linguistic experience.

Evidence for this effect is implied in the results of Withrow and of Furth and Pufall in investigations of learning of sequential material.[2] What is at issue in the present study is not sequential learning, *per se*, but the ability to discern information contained in sequentially presented material as it relates to the formation of rules of order.

Smith has shown that hearing *Ss* discover and use rules of order implicit in pre-

---

* Received for publication January 12, 1967. This research was supported, in part, by Research Grant No. RD–1479 from the Vocational Rehabilitation Administration. The authors are indebted to Dr. Lloyd Graunke and Warren Flower of the Tennessee School for the Deaf; John Nace and Peter Owsley of the Pennsylvania School for the Deaf; W. C. Yates, Carl Owen, Mrs. Lucille Britt, W. W. Wallace, Charles Shipley, and Mrs. Pearl English of the Williamson County, Tennessee Schools. Eva Neumann and Linda Maggart are due our sincere thanks for assistance in the collection of data.

[1] The authors' conclusions drawn from Elizabeth A. McCall's, A generative grammar of Sign. Unpublished Master's thesis, University of Iowa, 1965.

[2] Frank Withrow, Immediate recall by aphasic, deaf, and normally hearing children for visual forms presented simultaneously or sequentially in time. Unpublished Doctoral dissertation, Washington University, St. Louis, 1963; H. G. Furth and P. B. Pufall, Visual and auditory sequence learning in hearing-impaired children, *J. Speech Hearing Res.*, 9, 1966, 449–499.

structured letter pairs.[3] The present study is an attempt to determine whether deaf *S*s can also infer such rules. Two types of material were employed, nonsense forms (*PN*) as well as letters (*PL*), to determine whether the acquisition of rules may be a general disposition in deaf as well as in hearing *S*s. In addition, the relative effects of spatial and temporal order cues were studied by presenting the pair members successively (*SUC*), as well as simultaneously (*SIM*).

The task was similar to that used by Smith; the material to be learned by the *S*s is represented by the 5 × 5 matrix shown in Fig. 1.[4] The *S*s are shown, and told to remember, a list of the 20 individual letter pairs which are written in the cells of the matrix. Smith's *S*s, when asked to recall the pairs they had been presented, 'recalled' the five letter pairs of the diagonal (systematic errors), which they had not been given, significantly above chance expectation. Smith concluded that the *S*s had learned a rule rather than individual pairs.

*Method: (1) Subjects. S*s were drawn from three populations: (a) a profoundly deaf group composed of 80 *S*s from the Pennsylvania School for the Deaf and 48 from the Tennessee School for the Deaf, with an equal number from each group distributed over the four conditions; (b) an equal number (128) of fifth grade hearing *S*s, matched with the deaf on reading achievement; and (c) the same number of twelfth grade hearing *S*s, matched with the deaf on age. All of the *S*s had an *IQ* score of 80 or above.[5] Thirty-two *S*s from the three groups were randomly assigned to each condition of the experiment. Table I shows the characteristics of the three groups after they were divided into the four conditions.

		LAST LETTER OF PAIR				
		G	K	L	M	F
INITIAL LETTER OF PAIR	V	**	VK	VL	VM	VF
	H	HG	**	HL	HM	HF
	R	RG	RK	**	RM	RF
	X	XG	XK	XL	**	XF
	D	DG	CK	DL	DM	**

FIG. 1. MATRIX OF LETTER PAIRS USED IN SMITH'S EXPERIMENTS AND TWO CONDITIONS OF PRESENT EXPERIMENT

---

[3] K. H. Smith, Recall of paired verbal units under various conditions of organization. Unpublished Doctoral dissertation, University of Minnesota, 1963; K. H. Smith, Grammatical intrusions in the free recall of redundant strings of letters. Paper read at Midwestern Psychological Association meetings, St. Louis, May, 1964.

[4] Smith, Mediation and position learning in the recall of structured letter pairs, *Psychon. Sci.*, 2, 1965, 293; Grammatical intrusions in the free recall of structured letter pairs, *J. verb. Learn. verb. Behav.*, 5, 1966, 447.

[5] The reading achievement and *IQ* scores for the hearing *S*s were from the Metropolitan Achievement Test and the Lorge Thorndike *IQ* Test. The reading achievement scores for the deaf came from the advanced battery of The Stanford Achievement Test. Their *IQ* scores came from a variety of tests, predominatly the *WISC* and *WAIS*.

Materials. The pairs of letters shown in Fig. 1 were used in the *PL* conditions. A set of 10 nonsense forms was chosen from those constructed by Gibson.[6] Each nonsense form was assigned a position in the left or upper margin of the matrix, and 20 pairs of nonsense forms were generated by completing the non-diagonal cells of the matrix (*PN* conditions).

A recognitive task was employed to measure learning. A response-booklet contained 40 pairs listed down the side of the pages with a *YES* and *NO* printed beside each one. There were four forms of the booklet, differing only in the order the 40 pairs were listed. An equal number of each form was distributed to every group. The 40 response-pairs consisted of the 20 pairs actually presented to the *Ss*, plus the five systematic intrusions, five pairs composed of only initial members, five pairs made up of last members only, and five pairs with member order reversed. The 10 pairs composed of initial members only or last members only were included as filler items and were not counted in the scoring.

TABLE I

CHARACTERISTICS OF DEAF AND HEARING *S*s IN FOUR CONDITIONS

Condition and Group	Mean age	Mean *IQ*	Mean *R.A.**	Median *H.L.* (db)†
PL Simultaneous				
Deaf	18.2	97.0	4.4	93
5th grade hearing	10.2	102.6	5.5	—
12th grade hearing	17.3	101.7	11.6	—
PL Successive				
Deaf	17.2	99.0	4.5	85
5th grade hearing	11.1	103.1	5.5	—
12th grade hearing	16.7	108.9	12.1	—
PN Simultaneous				
Deaf	17.6	102.4	4.2	96
5th grade hearing	10.3	97.5	5.1	—
12th grade hearing	16.3	107.7	12.0	—
PN Successive				
Deaf	17.5	98.3	4.4	80
5th grade hearing	10.3	101.6	5.2	—
12th grade hearing	16.9	100.1	10.1	—

* Mean reading achievement grade equivalent.
† Hearing loss in best ear.

Procedure. The *Ss* participated in the experiment in groups ranging in size from 10 to 12. Response-booklets and instruction-sheets were passed out at the beginning of the experiment. Pennsylvania Deaf *Ss* and hearing *Ss* were told to read along with *E* as the instructions were read aloud. The instructions were given in manual language to the *Ss* at the Tennessee School for the Deaf.[7] Instructions included a detailed account of how the material was to be presented, and a sample pair was shown on the screen. The *Ss* were then told to remember as many of the pairs as they could. They were given an opportunity to ask questions about the procedure, and the session did not begin until *E* was satisfied that everyone understood the instructions.

---

[6] E. J. Gibson, Development of perception: Discrimination of depth compared with discrimination of abstract forms, in J. C. Wright and Jerome Kagan (eds.), *Monograph of the Society for Research in Child Development*, 28, 1963, 17.

[7] Different instructional channels were used at the two deaf schools because the Pennsylvania School for the Deaf requires speech reading in the classroom, while the Tennessee School for the Deaf does not.

The pairs were presented on slides by a Kodak Carousel projector, automatically paced by a Flexipulse timer. For the *SIM* conditions, the slides (which showed both pair members at once) were shown for 2 sec. each, followed by a 2-sec. pause. In the *SUC* conditions, the first slide (containing only the first member of a pair) was shown for 1 sec., followed immediately by the second slide for 1 sec. After both members of a pair in the *SUC* condition had been presented, there was a 2-sec. delay before the first member of the next pair was shown. The *Ss* were shown four randomizations of the 20 pairs. These randomizations were the same across all conditions.

After all the pairs were presented, the *Ss* were given 5 min. to complete the task of recognition. They were told to circle *YES* if they had just seen a particular pair or *NO* if they had not.

Results and discussion. Since the *Ss* were not told how many *YES* responses they could make and it is conceivable that some of them might circle *YES* after most or all of the pairs, the first analysis examined the total number of *YES* responses. (None of the *Ss*, in fact, circled all *YES*es.) The *PN* groups made significantly more *YES* responses than the *PL* groups ($F\ [1,372] = 44.1$, $p < 0.001$), and this difference appeared only in the means for the fifth and twelfth grade hearing *Ss*. The next analyses treated the distribution of *YES* responses over the different types of pairs.

TABLE II

MEAN NUMBERS OF PAIRS CORRECTLY IDENTIFIED

Group	*PN SIM*	*PN SUC*	*PL SIM*	*PL SUC*	Total
Deaf	15.4	14.6	14.5	15.0	59.5
5th graders	15.1	14.8	13.6	12.7	56.2
12th graders	16.6	15.0	16.1	15.0	62.7
Total	47.1	44.4	44.2	42.7	

The mean numbers of correct pairs (presented pairs) recognized by the *Ss* are shown in Table II. An analysis of variance on the number of pairs correctly recognized produced a significant groups-effect ($F\ [2,372] = 11.12$, $p < 0.001$), a significant materials-effect ($F\ [1,372] = 7.03$, $p < 0.01$), and a significant presentation-time effect ($F\ [1,372] = 5.92$, $p < 0.02$). Within the group-factor, the fifth graders recognized the fewest number of pairs, the twelfth graders recognized the most, and the deaf were in between. All differences between groups were significant ($p < 0.05$) as determined by a Newman-Keuls multiple comparison-test. A significant groups Z materials interaction (F [2,372] = $p < 0.05$) is attributable to the hearing *Ss*' recognition of more correct nonsense pairs than correct letter pairs. The greater number of *YES* responses made by the hearing *Ss* in the *PN* conditions was due to better recognition of the correct pairs in those conditions. This latter finding was not surprising, since letters are subject to interference-effects which new, discriminable nonsense forms do not entail.

A comparison of the types of errors by groups showed that each of the three groups recognized the systematic intrusions significantly more frequently than the other types of intrusions ($p < 0.05$, Newman-Keuls). There was also a significant difference between groups in the number of systematic intrusions, with the deaf recognizing fewer than the members of the fifth or twelfth grades. Evidently, then, deaf *Ss* did formulate rules in the learning of the material, but not to the extent that the hearing *Ss* did.

The most legitimate test for rule-inference within the conditions is a comparison between systematic errors and reversal errors. This relationship for the deaf is shown in Fig. 2. Apparently, the optimal condition (as shown by more systematic errors and fewer reversal errors) was the simultaneous presentation of nonsense pairs. Of the 32 deaf *Ss* serving in that condition, 28 (88%) recognized more systematic errors than reversal errors. On the other hand, the deaf's worst performance was in the *PN SUC* condition. (The hearing *Ss* made more systematic errors on nonsense pairs than on letter pairs, but they also made more reversal errors on nonsense pairs.)

Several alternative explanations can be offered to account for this differential performance by the deaf. If their inferior performance on the successively presented nonsense forms is due to an inability to transfer a first-second type of learning (successively presented pair members) to a left-right recognitive task, one might surmise that this results from a lack of correspondence between written English and sign language. Since hearing *Ss* experience a one-to-one correspondence between the language that they speak and hear and the language as they read it (in terms of word order), it is reasonable to assume that translating from a temporal first-second sequence to a spatial left-right arrangement is quite natural and automatic. The deaf who use sign language do not have this advantage.

If, however, the deaf's poor performance on the successively presented nonsense forms is due to an inability to distinguish or remember positions of the items regarding first-lastness, the explanation must lie in some other process. There is still not enough known about either cross-modal transfer of different types of learning or 'what is learned' when children learn the rules of language to attempt more than a very tentative explanation.

It may be that learning of temporal position (as postulated by Braine and Smith)[8] is a type of learning that is more common with auditory than with visual material. The temporal ordering of visual inputs may be facilitated by past experience with a language that requires strict adherence to rules of order, the absence of which would impair performance on such sequences for the deaf. It is interesting to note that the deaf and hearing *Ss* performed comparably on the *PL* task under both *SIM* and *SUC* conditions. The more typical performance of the deaf on these tasks may be due to the fact that finger spelling of the pairs provided for both symbolic coding and temporal ordering of the items.

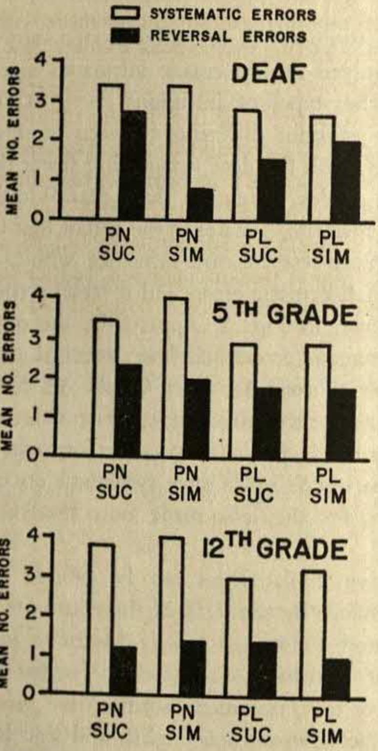


FIG. 2. MEAN NUMBERS OF SYSTEMATIC INTRUSION AND REVERSAL ERRORS MADE BY THE *Ss*

---

[8] M. D. S. Braine, On learning the grammatical order of words, *Psychol. Rev.*, 70, 1963, 323–348; On the basis of phrase structure: A reply to Bever, Fodor, and Weksel, *Psychol. Rev.*, 72, 1965, 483–497; K. H. Smith, Grammatical intrusions in the recall of structured letter pairs: Mediated transfer or position learning? *J. exp. Psychol.*, 72, 1966, 580–588.

SUMMARY

The present study was designed to determine whether performance by deaf *S*s on learning rules of order was comparable to that of hearing *S*s. The material presented to *S*s was 20 individual letter-pairs or nonsense form pairs. These were constructed by designating five letters or nonsense forms as initial pair members and five letters or nonsense forms as last members and then combining every initial member with every last member to form 25 pairs. Five of these pairs were not shown to the *S*s. Half the *S*s were shown both members of a pair simultaneously and the other half were shown one member of a pair and then by the second member of the pair. Learning was measured by a recognitive task. The tendency to recognize the five omitted pairs was used as a measure of rule-inference.

Over-all, the deaf recognized fewer of the omitted pairs than the hearing *S*s, indicating less tendency to form rules about the organization of the material. The deaf's worst performance was on the successively presented nonsense forms. Several hypotheses are proposed to account for this finding.

A COMPARISON OF RECALL AND BINARY RECOGNITION

By Bruce A. Pappas, K. Lynne Hepburn, and Milton D. Suboski, Queen's University
Kingston, Ontario

Recognition has traditionally been regarded as a more sensitive measure of retention than recall.[1] Recent experiments have suggested, however, that when the number of alternatives on the recognition- and recall-tests[2] and the similarity between correct and incorrect alternatives on the recognition-test are equated,[3] the apparently greater sensitivity of recognition is reduced, if not eliminated. Furthermore, the temporal sequence of recognition and recall appears to modify the effect that these two methods of measuring retention have upon each other. Recognition has been found to be poorer after recall than before recall, while on the other hand, recall is better after recognition than before.[4] This difference was attributed by the authors to the dropping out of weak associations during an initial recall test, thus lowering subsequent recognition performance, and to the rehearsal function of the recognition test which serves as a final learning trial for later recall.[5]

Tests of recognition in all of the studies cited above have been of the discriminative variety, with the *S*s required to select the correct from among incorrect alternatives. Discrimination of the correct response from a display of alternatives would generally be expected a somewhat easier task than 'absolute' recall which requires total reconstruction of the item. Furthermore, the *S*s would seem to benefit more from partial learning[6] on this test of recognition than on recall. On the other hand, a binary 'yes-no' test with

---

* Received for publication February 27, 1967. This research was supported by grants from The Society of the Sigma Xi and the Queen's University Arts Research Committee. The authors are grateful to Mrs. Pat Pappas for her assistance in the experiment.

[1] C. W. Luh, The conditions of retention, *Psychol. Monogr.*, 31, 1922 (No. 142) 1–87.

[2] R. Davis, N. S. Sutherland, and B. R. Judd, Information content in recognition and recall, *J. exp. Psychol.*, 61, 1961, 422–429.

[3] J. A. McNulty, An analysis of recall and recognition processes in verbal learning, *J. verbal Learn. verbal Behav.*, 4, 1965, 430–436.

[4] Leo Postman, W. O. Jenkins, and D. L. Postman, An experimental comparison of active recall and recognition, this JOURNAL, 61, 1948, 511–519.

[5] Postman, Jenkins, and Postman, *op. cit.*, 51.

[6] J. A. McNulty, A partial learning model of recognition memory, *Canad. J. Psychol.*, 20, 1966, 302–315.

simple test-items would eliminate partial learning aids and to a degree, any benefit due to the method of display. As a result, the superiority of both performance and rehearsal function of the recognition-test should be minimized.

In the present experiment, a modified *RTTT* design was used with the second test (T_2) either a binary recognition or a recall trial. Learning materials were simple letter-digit paired-associates. Besides providing for comparison of recall and binary recognition-test accuracy, any differences in the rehearsal-function in recall and recognition should be reflected by the retention-function displayed on the third test-trial. A central difficulty in comparing recognition and recall has been selecting a measure equally sensitive to variations in performance. Response proportions alone provide only a limited binary categorization, while the applications of guessing corrections require the assumption of thresholds for correct and incorrect response strengths.[7] Recent applications of signal detection-theory to recognition[8] and recall[9] have suggested that this method provides both sensitive and possibly comparable assessments of performance for both types of retention measure. In the experiment reported here, confidence ratings were used to construct operating characteristics (*ROC*s) for assessing accuracy of recall and recognition.

Method and procedure. The *S*s were nine men and nine women, volunteers from an introductory course in psychology.

The learning materials were 10 lists, each consisting of a random pairing of the letters A–J as stimulus-items and the numbers 0–9 as responses. The S–R pairs were placed on 3 × 5-in. cards and presented at an exposure rate of 4-sec. per pair by hand.

On the recall-tests, the *S*s worked at the rate of 7-sec. per page through a 10 page booklet. Each page contained a stimulus-letter, a space to write in the response-number, and a 5-cm. 'confidence'-line which had the left end labelled 'guess' and the right end labelled 'sure'. The *S*s were instructed to draw a vertical line across the 'confidence'-line to indicate their degree of confidence in the correctness of their responses. On the recognition-tests, the *S*s worked through a similar booklet except that responses were already inserted. They were instructed to mark the pairing as correct or incorrect and to rate their confidence in their decision.

The procedure followed an *RTTT* paradigm with T_2 being either a recall-test or a recognition-test, while T_1 and T_3 were always recall-tests. Order of the recall- and recognition-series was random among lists and across the *S*s. On the five recognition-tests for each *S*, a random pairing of from three to seven correct pairings, with an average of five correct, were interspersed with incorrect pairs. The *S*s were informed that across lists, half of the pairs would be correct.

---

[7] J. P. Egan, Recognition memory and the operating characteristic, *USAF Operat. Appl. Lab. tech. Note,* (No. 51–58), 1958, 3.

[8] B. B. Murdock, Jr., Signal-detection theory and short-term memory, *J. exp. Psychol.,* 70, 1965, 443–447.

[9] B. B. Murdock, Jr., The criterion problem in short-term memory, *J. exp. Psychol.,* 72, 1966, 317–324.

Order of items within booklets was random across tests, lists, and *S*s. The first test was administered immediately after list presentation, while 15-sec. intervals elapsed between subsequent tests.

Results. (1) Response-proportions. Response-proportions for recognition- and recall-lists, pooled over *S*s and lists, are shown in Fig. 1. For items correct on T_1-recall, correct recall on T_2 increases the proportion of correct T_3-recalls over that for items correctly recognized (hits and correct rejections) on T_2 by 0.15. Correspondingly, for items incorrect on T_1, *correct* T_2-recall in comparison to recognition again facilitates subsequent recall, the difference in proportion being 0.26. Incorrect T_2-recall, in comparison with incorrect recognition (miss or false alarm) decreases the proportion of T_3 corrects for both T_1 corrects and incorrects.

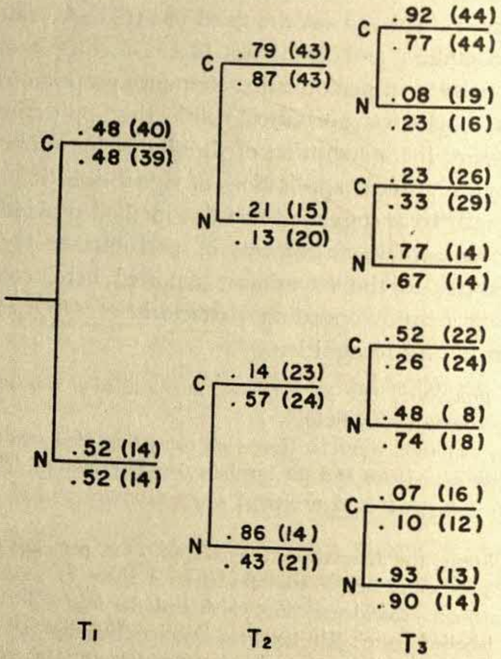


FIG. 1. PROPORTIONS OF CORRECT AND INCORRECT OVER TRIALS

Data for recall-series are shown above the line; mean confidence ratings are shown in parentheses.

Over-all proportions of correct responses across the three trials were

0.48, 0.44, 0.43 for the recall-series, and 0.48, 0.72 and 0.44 for the recognition-series.

The proportion of shifts to a new response on T_3 after a correct T_2-recall was 0.27, compared with the proportion of 0.37 after recognizing a correct pair. Correspondingly, the proportion of shifts after an incorrect recall was 0.68 compared with the 0.75 after a false alarm.

(2) Confidence-ratings and ROCs. Fig. 1 also displays mean confidence-ratings for sequences of correct and incorrect over trials. The *S*s appear to have rated correct recalls and correct recognitions equivalently for both T_1

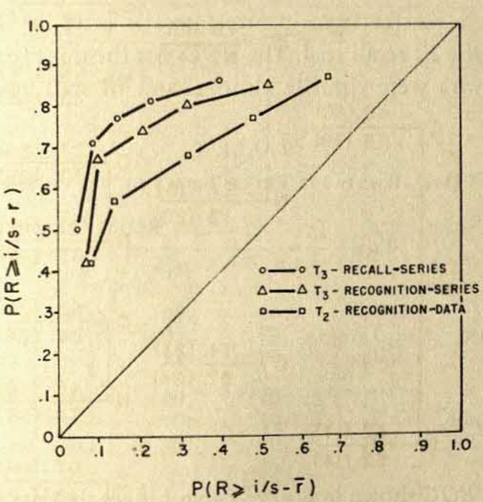


FIG. 2. *ROC*S FOR RECALL AFTER RECOGNITION AND RECALL-TRIALS AND FOR RECOGNITION-TRIAL

corrects and incorrects, as indicated by sign-tests performed on individual ratings ($p > 0.40$). Incorrect recognitions are, on the other hand, rated more highly than incorrect recalls for both T_1 corrects and incorrects (*p*s < 0.05 by sign-tests). The T_1 and T_3 ratings clearly correspond with the classifications of correct and incorrect.

To construct the *ROC*s, the confidence-line was first divided into six segments of decreasing confidence. Category "one" included all responses rated 50. The next 'five' categories included respectively responses rated 49 to 40, 39 to 30, 29 to 20, 19 to 10, and 9 to 0. Both recognition- and recall-curves are of the Type II or response conditional variety.[10] For recall-data, the proportion of correct recalls (S—R) assigned confidence Category or higher ($R > i$) were pooled across the Ss for the six categories and plotted against the corresponding proportions for incorrect recalls. A similar technique yielded recognition-curves by treating hits and correct rejections as correct responses and false alarms and misses as incorrect.

---

[10] F. R. Clarke, T. G. Birdsall, and W. P. Tanner, Jr., Two types of ROC curves and definitions of parameters, *J. acoust. Soc. Amer.*, 31, 1959, 629–630.

Fig. 2 displays two T_3 recall *ROC*s, the uppermost curve showing T_3 data for the recall series, and the curve immediately below it that for the recognition-series. The lowermost *ROC* is for the recognition-trial. Qualitative assessment of the recall curves in terms of their relative displacement from the chance diagonal indicates that the Ss were less accurate in discriminating their correct from incorrect recalls after the T_2 recognition-trial than after the T_2 recall trial. The *ROC*s for the three trials of the recall series (not shown) were virtually identical and are well represented by the T_3 recall series *ROC* shown in Fig. 2. Table I gives the data points for these curves. Apparently, there was no change across recall trials in the Ss' ability to discriminate correct from incorrect recalls. Comparing the *ROC* for the recognition with the T_3 recall curve indicates that the Ss were considerably less accurate in discriminating between correct and incorrect recognitions than recalls.

TABLE I

ROC DATA-POINTS FOR THREE TRIALS OF RECALL-SERIES

	Trial	Rating Category				
		R_1	R_2	R_3	R_4	R_5
Correct response	T_1	.48	.70	.78	.84	.89
	T_2	.47	.69	.78	.84	.89
	T_3	.50	.71	.77	.82	.87
Incorrect response	T_1	.05	.09	.17	.29	.47
	T_2	.05	.09	.16	.28	.43
	T_3	.05	.09	.16	.24	.44

Figure 3 shows T_2 *ROC*s for recognition and recall as a function of correct or incorrect recall on T_1. To construct these curves, cumulative conditional proportions for items correct on T_2 given either correct or incorrect on T_1 were plotted against the corresponding conditional proportions for items incorrect on T_2 irrespective of their status on T_1. Inspection of the curves indicates that after a prior correct recall, recall performance is consistently better than recognition performance. On the other hand, after an incorrect recall no consistent difference is apparent. Discrimination of correct from incorrect responses for both recall and recognition is, however, above chance after an incorrect recall.

Discussion. Although the response-proportion data indicate no differences in the over-all proportion of responses correct on T_3 between the recall- and recognition-series, the greater probability of recalling a response correctly after its correct recall than after its correct recognition suggests stronger reinforcing properties for recall than for binary recognition. The slightly greater proportions of T_3 corrects after incorrect recognitions than after incorrect recalls and the response shift data also suggest that recognitions are less likely to be repeated than recalls. Furthermore, the T_3 *ROC*s similarly indicate that recalled responses are less discriminable as correct or incorrect after binary recognition.

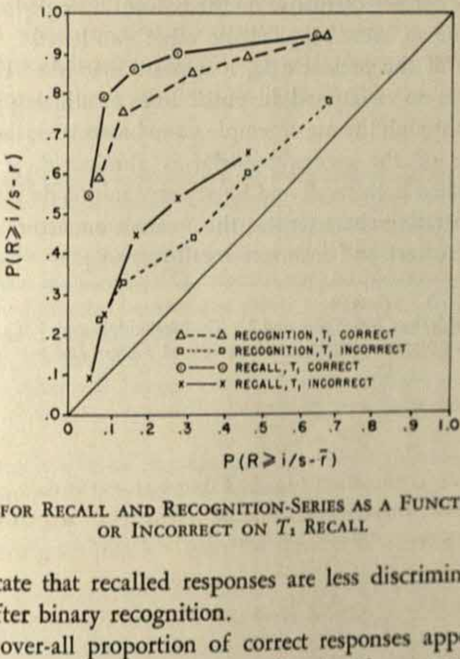


FIG. 3. T_3*ROC*s FOR RECALL AND RECOGNITION-SERIES AS A FUNCTION OF CORRECT OR INCORRECT ON T_1 RECALL

While the over-all proportion of correct responses appears to decline slightly over the three trials of the recall-series, the corresponding *ROC*s show no change. Apparent learning on test-trials has been observed by Jones[11] who interpreted this result as contradictory to the *all-or-none* retention process as proposed by Estes.[12] Analysis of the *ROC*s in the present study suggests that discriminability of both recall and recognition response is above chance after a prior incorrect recall (a result which is also in opposition to the *all-or-none* hypothesis), although there is no over-all improvement over test-trials. Furthermore, the recognition-test does not appear to be more sensitive to these weak associations than recall. The dissimilarity between the present results and those of Postman, who found recall to be facilitated by a prior recognition-trial, may be explained by the difference in items and recognition tests.[13] In the test used by Postman, *S*s were required to select the correct three letter nonsense syllable from one incorrect alternative differing by one letter, and two other alternatives with all three letters different from the original. The *S*s could thus discriminate the correct alternative on the basis of a single letter and having successfully done so, were supplied the other two letters.

---

[11] J. E. Jones, All-or-none versus incremental learning, *Psychol. Rev.*, 69, 1962, 156–160.

[12] W. K. Estes, B. L. Hopkins, and E. J. Crothers, All-or-none and conservation effects in the learning and retention of paired-associates, *J. exp. Psychol.*, 60, 1960, 329–339.

On the basis of the present data, it appears that when the recognition-task is binary, yes-no variety and the possibilities of utilizing partial learning are minimized through the use of simple paired-associates, the superiority in rehearsal value of the recognition-trial is eliminated, if not reversed. Furthermore, when both recall and binary recognition-data are analyzed in terms of the operating characteristic, the *S*s are more accurate in discriminating between correct and incorrect recall-responses.[14]

---

[13] Postman, *op. cit.*, 511–519.

[14] A similar result has been reported by D. McFadden and J. G. Greeno, Verbal learning: a some-or-none process, *mimeographed paper*, Indiana University, 1966

FEEDBACK AND ACCURACY OF TARGET POSITIONING IN A HOMOGENEOUS VISUAL FIELD

By MILTON S. KATZ, Washington, D. C.

The ability of subjects (*Ss*) in a homogeneous visual environment to position targets displaced from geometric center along a horizontal axis was studied by Miller and Hall.[1] Their *Ss* adjusted a target subtending approximately 25′ of visual angle to a given field-center or reference-point with an accuracy of about ±3° and a high degree of consistency in the absence of structural visual detail.

Katz, Metlay, and Cirincione confirmed the above findings, using a target of similar size.[2] When, however, a smaller 8′ target was used, errors in placement exceeded ±3° in most cases and, in addition, variability of placements increased markedly.

Since these works in homogeneous environments investigated horizontal accuracy of target acquisition and position adjustment, a study by Metlay *et al.* was conducted to define possible differences between the ability to position a target in the vertical and horizontal axes.[3] Although firm conclusions could not be based on the vertical accuracy findings, the data did suggest that feedback, *i.e.*, repeated indications of where geometric center was located, could affect the accuracy of positioning in a homogeneous field. One *S* consistently produced no error in vertical and horizontal target placements.

The present report is of an investigation of the extent to which the accuracy of such target placements is influenced by feedback. The hypothesis under test is that if an *S* is provided with information about the magnitude of his errors from the geometric center, his performance will improve in both constant error-score and variability.

Method: (1) Apparatus. The basic elements of the experimental room are shown in Fig. 1. To approximate a homogeneous visual environment, a 270° curved, flat-white projection screen with an internal diameter of 20 ft. was used. The screen was constructed of fiberglass bonded to a frame of fiber-board and steel. To complete the homogeneous surround, the floor and *S*'s body were covered with white fabric

---

[*] Received for publication January 5, 1967. This work was conducted at the U. S. Naval Training Device Center, Port Washington, New York.

[1] J. W. Miller and R. J. Hall, Accuracy of orientation and positioning in homogeneous visual fields, *Aerospace Med.*, 34, 1963, 337–341.

[2] M. S. Katz, W. Metlay, and P. A. Cirincione, Effects of stimulus and field size on the accuracy of orientation in the homogeneous environment, *Percep. mot. Skills*, 20, 1965, 167–172.

[3] W. Metlay, M. S. Katz, P. A. Cirincione, and G. C. Tolhurst, Vertical orientation in a homogeneous environment, U. S. Naval Training Device Center, *Technical Report IH-19, 1964.*

The field was illuminated to a brightness of 107.6 mL. by a series of 200-w. incandescent G.E. bulbs, arranged, as shown in Fig. 1, in four banks of five bulbs each. Departure from homogeneity, as measured by the Spectra Pritchard Photometer did not exceed 1 mL.

The target to be positioned was a circular spot of light subtending 8′ of visual angle at the eye. Total luminance of the superimposed target was 136 mL. The spot was produced by a 500-w., 35-mm. Revere Slide Projector mounted on a tripod which was adjustable in the horizontal and vertical axes. The projector on the tripod, as shown in Fig. 1, was located at a distance of 12 ft. 7 in. from the base of the screen, and was so positioned that its forward throw directly bisected the field.

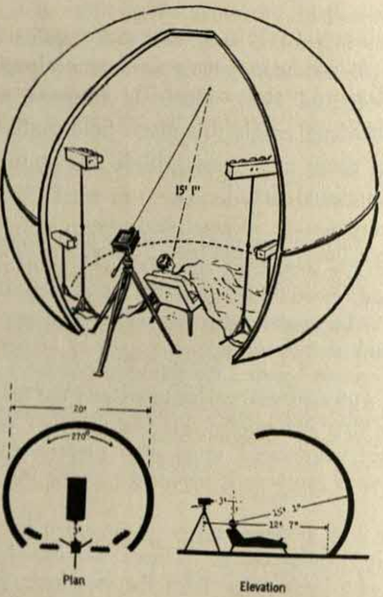


FIG. 1. ILLUSTRATION OF EXPERIMENTAL ROOM AND APPARATUS

S reclined on a chaise longue 3 ft. in front of the projector and 3 ft. below it, so that his forward gaze bisected the screen as well. A head-restraining device was not employed.

The projector, taken to read 0° within an accuracy of 0.5° of arc, was so positioned upward in the vertical plane that the angle of elevation of the target was 41°. The *S*'s position placed the target on his line of regard.

Procedure. Before the start of an experimental session, *S* was shown the horizontal path the target would travel. He was instructed that a target would appear somewhere along this path, and that his task was to search for the target, and, after locating it, reposition it to the center of the field by telling *E* which way to move the target; when to stop; and the direction and extent of minor adjustments for centering.

In each experimental session, targets were positioned from one of two field locations, either 23° right or 23° left. The *S* was able to keep his eyes open at all times, since *E* occluded the projector while positioning the target to one of the initial field locations.

Four men, three experimental and one control *S*—ages 25 (*WM*), 27 (*PC*), 38 (*MK*), 27 (*JC*) yr. old were employed. The treatments listed below were administered to each *S* in the following order:

(1) Baseline performance. Forty trials of target placement to establish a performance baseline.

(2) Limit, ±2.5°. Knowledge of results was given *S* following each target placement. If placement was in error, *E* said "No" and repositioned the target to the geometric center while *S* watched. If the placement was within the limits set, *E* said "all right" and proceeded to the next trial. The limits of error were set at ±2.5°, and this session continued until *S* reached the criterion of no more than one error in 20 consecutive trials.

(3) Limit, ±1.5°. A second session, identical to that above in all respects except that the acceptable limits of error were reduced to ±1.5°.

(4) Limit, 0°. A third session, identical to the two preceding in all respects except that the acceptable limits of error were reduced to 0°, *i.e.* no error was allowed on any trial. Furthermore, this session terminated after *S* had received 60 trials.

(5) Immediate retest. On the day following the completion of treatment (4) 40 trials were administered to each S. In this session, no knowledge of results was given.

(6) Delayed retest. A second retest of 40 trials without knowledge of results was given to each *S*, one week after immediate retest.

The control *S* received the same number of sessions, separated by the same amount of time. This *S* received no feedback in any of his first six sessions, which consisted of 40 trials each.

(7) Limit, 0°. The control *S* was given an additional session which duplicated Treatment (4) (above).

Results. Mean constant error (subjective center) in degrees, and standard deviation, as a function of experimental treatment, are presented in Figs. 2 and 3 respectively. In Fig. 2 constant errors in the Baseline session were −2.4, −2.7, −6.7, and −4.8° for *Ss*: *PC, MK, WM,* and *JC* (control) respectively; (minus signs indicate that errors were located in the left half of the field). Variability about these subjective centers (Fig. 3) were 2.1, 1.5, 1.7, and 2.2 SD.

In the first experimental session (2.5° Limit), all experimental *Ss* met criterion (*PC* and *MK* in 20 trials, and *WM* in 42) as shown by the decrease in error scores in Fig. 2. A slight increase in average variability occurred. The performance of the control *S* shows an increase in constant error from −4.83° to −5.70°, and a decrease in variability from 2.5 to 1.37 SD.

In the second experimental session (Limits reduced to ±1.5°), all experimental *Ss* were again able to meet the criterion (*MK* in 20 trials, *PC* in 22, and *WM* in 31 trials). For the experimental *Ss*, constant error was further reduced slightly while variability decreased. In the case of the control *S*, all performance measures increased slightly.

In the final experimental session (Limit 0°), with any deviation from geometric

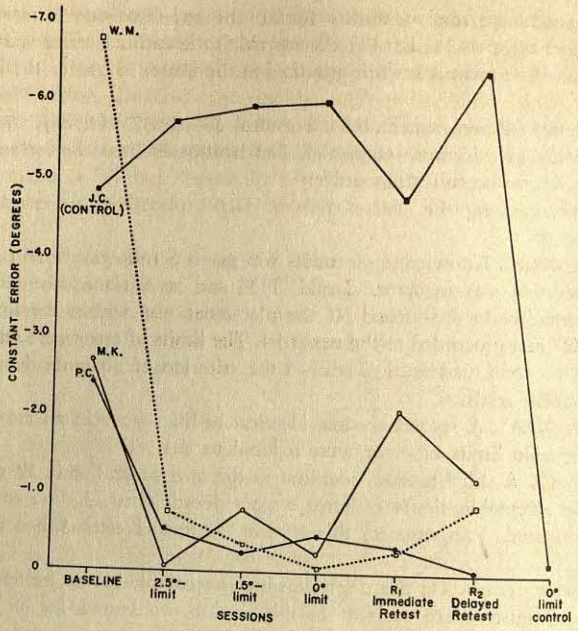


FIG. 2. CONSTANT ERROR, OR SUBJECTIVE CENTER, AS A FUNCTION OF EXPERIMENTAL TREATMENT

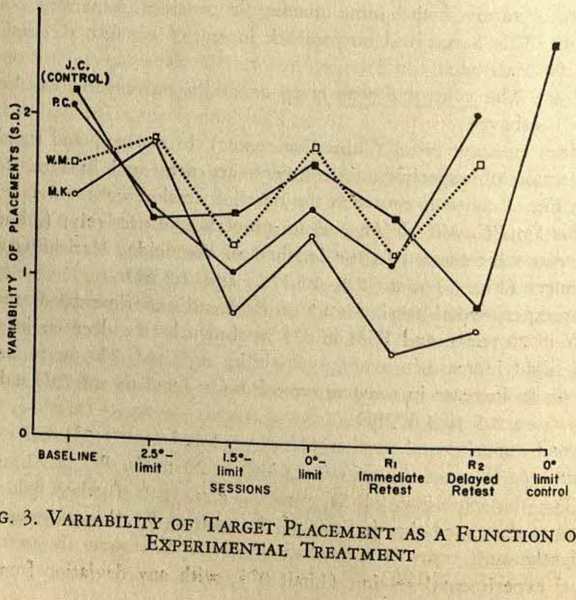


FIG. 3. VARIABILITY OF TARGET PLACEMENT AS A FUNCTION OF EXPERIMENTAL TREATMENT

center considered in error, variability for all the experimental *Ss* increased. The average constant error decreased. For the control *S*, the constant error was essentially the same as in the previous session although SD increased slightly.

Immediate retest, conducted without informational feedback, resulted in a decrease in variability for all experimental *Ss*. For *MK* and *WM*, variability was actually smaller than in any previous session. *PC*'s variability equaled the lowest score obtained by him previously. Constant error increased slightly. Constant error and variability decreased for the control *S* from −6.0° to −4.8°, and from 1.7 to 1.4 SD.

On delayed retest, constant error scores for the experimental *Ss* remained essentially unchanged while variability increased. In the case of the control *S*, constant error increased from −4.8° to −6.4°, while variability decreased from 1.4 to .9 SD. With 0° Limit training, the control *S*'s constant error showed a marked drop from −6.39° to −.13°. His variability, however, exceeded baseline level, and was the highest of all sessions.

Discussion. The data of the control *S* and one of the experimental *Ss* (*WM*) confirm the finding by Katz, *et al.* that an 8′ target may elicit constant errors in excess of 3.0° (Fig. 2).[4] The two remaining experimental *Ss* (*MK, PC*) performed with constant errors slightly less than 3.0°, but greater than in those of the former study with targets larger than 8′. *Ss MK* and *PC* had considerably more prior experience positioning targets in the homogeneous field than *WM*, and the control, *JC*, was new at the task. The data of the experimental *Ss* are not, therefore, directly comparable to those of the control *S*.

It would be reasonable to predict that, after the control *S* had experienced numerous positioning trials, his error and variability scores would decrease. His constant error did not improve with practice (Fig. 2) and, in many cases, was greater than at baseline level. On the other hand, variability in positioning the target did decrease, and paralleled performance observed for the more experienced *Ss* (Fig. 3).

Relative to baseline and control performance, the experimental treatments (feedback of results) reduce constant error (Fig. 2) or bring subjective center closer to geometric center. The effect is substantial both during feedback trials and on retests, one day and one week later. Evidently, use can be made by observers of transient informative feedback to orient very accurately in a homogeneous field.

The experimental *Ss*' constant errors undergo no marked change among the three experimental training sessions, with informative feedback at 2.5, 1.5, and 0 degrees of accuracy. Annett's review of the literature on knowledge of results suggests that extreme specificity does not always result in greater learning efficiency, but may reduce efficiency in some cases.[5] Similarly, Bilodeau and Rosenbach report that informative feedback can often be refined beyond its capability to serve response accuracy.[6]

Standard deviations, or variable error, for both experimental and control *Ss* appear to improve gradually in the course of the six sessions. (Fig. 3). The change appears to be related to practice rather than to any effect of the feedback regimen on variability.

---

[4] Katz *et al*, *op. cit.*, 7.

[5] J. Annett, The role of knowledge of results in learning: A survey, U. S. Naval Training Device Center *Technical Report* 342–3, 1961.

[6] E. A. Bilodeau and J. H. Rosenbach, Acquisition of response proficiency as a function of rounding error in informative feedback, Human Resources Research Center *Res. Bull.*, 53–21, 1953.

The increase in the control *Ss'* standard deviation during the trials of the 0° Limit feedback following the main experiment is probably partly a consequence of being subjected without preparation to the most rigorous performance-requirement. The suggestion of a similar elevation of variable error for the experimental *Ss* also appears in Fig. 3. It seems likely from these data that whatever improvement in constant error may come from the explicit and extremely accurate feedback of the 0° Limit trials, it is paid for with increased variable error.

Conclusions. The feedback technique used in this study did not result in near-zero constant and variable errors such as were found with one *S* in Metlay *et al.*[7] Feedback appears to have been instrumental in reducing constant errors as predicted not only during experimental sessions, but in retest sessions when feedback was no longer available. Variable error, contrary to prediction, seems to have been responsive to mere practice, except for a possible deleterious effect of the extremely specific 0° Limit feedback.

Further experimental work is needed to enable the feedback technique to operate more efficiently, and to identify the limits within which additional increments of informative accuracy result in improved performance. The reversal in variable error level in the 0° Limit feedback suggests that too much information interferes with an important aspect of positioning performance and that a similar effect can be looked for in other tasks.

SUMMARY

Three *Ss* were provided with feedback about error magnitude following each attempt to position a target at the geometric center of an empty visual field. Upon achievement of a criterional level of accuracy, feedback was discontinued, and accuracy of performance was compared with that of a control *S*. Experimental feedback of results reduced constant error both during feedback and on subsequent retests without feedback. Variable error (standard deviation) was apparently responsive to practice, rather than to informational feedback, except that extremely accurate feedback seemed to produce a deleterious effect on performance.

---

[7] Metlay *et al., op. cit.,* 2–4.

STIMULUS-REPETITION AND COLOR DISCRIMINATION

By MICHAEL H. SIEGEL, Edgewood Arsenal, Maryland

It has become abundantly clear that a number of experimental variables can affect sensitivity to color differences. The choice of psychophysical method,[1] the availability of response categories,[2] the wavelength composition of the standard stimulus,[3] and the duration of the stimulus-exposure are but a few of the variables which have profound influences on color discrimination.[4] In the present report, we have investigated the effect of the order of presentation of stimuli upon the discrimination of color sensitivity.

In what has become standard practice for this laboratory, several variable stimuli are presented to an observer (*O*) in a random order. Would sensitivity-scores change if the order were made non-random?

Two reasons prompted us to consider the presentation of a non-random order. A random sequence requires frequent changes of instrument-settings. The leads to lengthy sessions which in turn fatigue the *O*s. If stimuli were presented in repeated blocks rather than at random, at least some of this problem would be solved. There is some indication from the literature that non-random sequences do not adversely affect discriminations of lifted weights or of certain auditory stimuli;[5] however, no direct evidence is presently available on how such a procedure influences color discrimination.

A second reason for adopting a non-random order was to assess a growing body of research purporting to measure various aspects of sensitivity by presentation of stimuli in blocks. An early and well-known example of this form of stimulus presen-

---

* Received for publication December 20, 1966. From the Edgewood Arsenal Research Laboratories.

[1] H. R. Blackwell, Studies of psychophysical methods for measuring visual thresholds, *J. opt. Soc. Amer.*, 42, 1952, 606-616; M. H. Siegel, Discrimination of color. I. Comparison of three psychophysical methods, *J. opt. Soc. Amer.*, 52, 1962, 1067–1070; The selection of judgment categories in color discrimination, *Psychon. Sci.*, 2, 1965, 151–152.

[3] M. H. Siegel and F. L. Dimmick, Discrimination of color: II. Sensitivity as a function of spectral wavelength, 510 to 630 mμ, *J. opt. Soc. Amer.*, 52, 1962, 1071–1074; M. H. Siegel, Discrimination of color: IV. Sensitivity as a function of spectral wavelength; 410 through 500 mμ, *J. opt. Soc. Amer.*, 54, 1964, 821–823; M. M. Connors and M. H. Siegel, Differential color sensitivity in the purple region, *J. opt. Soc. Amer.*, 54, 1964, 1374–1377.

[4] Siegel, Color discrimination as a function of exposure time, *J. opt. Soc. Amer.*, 55, 1965, 566–568.

[5] Dorothy J. Shaad and Harry Helson, Group presentation in the method of constant stimuli as a time-saving device, this JOURNAL, 43, 1931, 422–433; J. F. Jerger, On the independence of successive responses in the quantal psychophysical method, this JOURNAL, 68, 1955, 145–147.

tation is the neural quantum theory.[6] A review of the then current literature on this theory was performed a decade ago.[7] Although the theory is derived from studies in audition, some vision research has also been directed upon it.

The original theory was concerned with the sensory intensity-continuum. It was assumed that the neural structure involved in the perception of a sensory continuum are divided into functionally distinct units. A stimulus of a given magnitude excites, at a particular instant, a certain number of these quantal units, and in order for an increment to be noticeable, it must excite at least one additional quantum.

The theory predicts a discontinuous step-like increase in response-probability as the stimulus-value is changed. The more typical psychophysical curve, by contrast, shows a continuous increase in response probability as the stimulus is changed. Both types of curves are presented in Fig. 1.

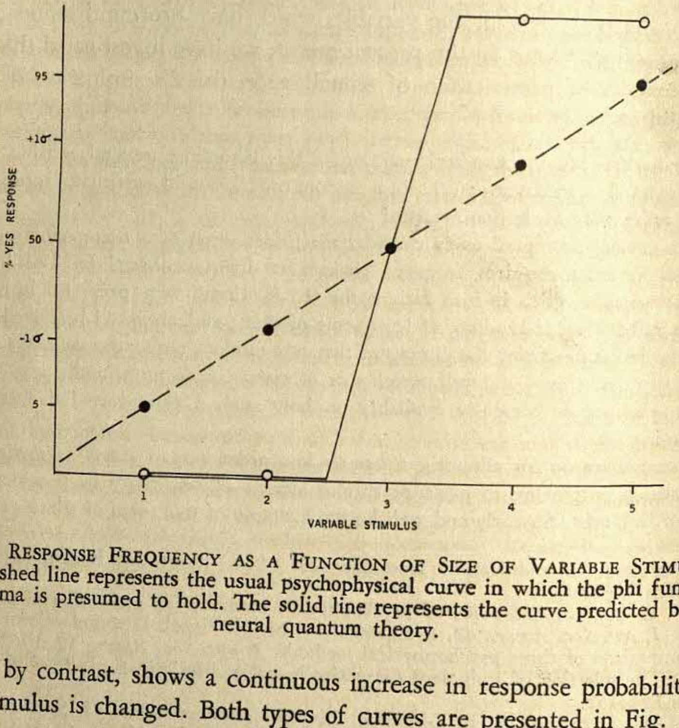


FIG. 1. RESPONSE FREQUENCY AS A FUNCTION OF SIZE OF VARIABLE STIMULUS
The dashed line represents the usual psychophysical curve in which the phi function of gamma is presumed to hold. The solid line represents the curve predicted by the neural quantum theory.

To detect the quantal increases in response, the experimenter (*E*) must depart from the usual psychophysical procedures in which stimuli appear in a random order. In this way, the quantal distribution will not be masked by shifts in *O*'s criterion. This departure from standard psychophysical procedure, necessary as it may be, is worthy of careful attention. Since the experimental evidence for the neural quantum theory rests upon data in which the same stimulus-value is presented 25 times in a row, it is of some significance to know whether or not the order of presentation affects the results.

---

[6] S. S. Stevens, C. T. Morgan, and John Volkmann, Theory of the neural quantum in the discrimination of loudness and pitch, this JOURNAL, 54, 1941, 315–335.

[7] J. F. Corso, The neural quantum theory of sensory discrimination. *Psychol. Bull.* *53,* 1956, 371–393.

Apparatus. A Farrand monochromator and a xenon arc produced the stimulus. A sector shutter with a clutch provided discrete 0.2-sec. exposures of the stimuli. The stimulus-field was circular and subtended two degrees at the *O*'s eye. The upper half served as the standard. It was set to a wavelength of 570 mμ, which most *O*'s call yellow-green, and a luminance of 0.2 ft.-L. Calibrations were made both before and after the experiment upon the wavelength setting of the monochromator, the luminance level of the stimulus and the duration of the exposure-time. No change in the pre-experimental levels could be detected.

Observers. Four members of the laboratory staff served as *O*'s. All received extensive training in making the required discrimination before data were collected. All were free from defects of color vision. The non-emmetropic *O*s wore corrective lenses and all observations were made with the right eye.

Procedure. Before each experimental session, a pretest was performed to allow the *O*s to equate the brightness of each of the variable stimuli to that of the standard and to determine the range of varible stimuli to be presented. For each experimental condition, the stimulus-range consisted of five steps each of which was presented a total of 10 times. At each session, two separate measurements were made: one, for wavelengths shorter than the standard, and the second, for wavelengths longer than the standard.

Experimental conditions. The first condition was a standard method of constant stimulus-differences. Fifty stimuli, consisting of 5 variable stimuli, were presented a total of 10 times each in a random order. In the second condition, each stimulus-difference was presented two times in a row before another, randomly selected, stimulus-difference was presented. In the third condition each stimulus was presented five times in a row. In the fourth condition, each stimulus-difference was presented 10 times in a row.

The order of appearance of each of the four conditions was randomized and then reversed for each *O*. The resulting order of these 8 sessions was then repeated twice. This permitted each *O* to experience each of the 4 conditions 6 times.

Results. Curves of frequency of positive responses were recorded on normal probability paper. The measure of sensitivity, the standard deviation, was defined as the difference in wavelength between the points at 50% and at approximately 84%.

Fig. 2 presents the experimental results for this study. The four points on the abscissa represent from left to right Conditions 1 through 4. The ordinate is size of standard deviation in millimicrons. The first and most obvious result is that large individual differences occur among the four *O*s. There is no apparent relation between the sensitivity scores obtained in the first condition and the scores for the other conditions. There is, however, a relation among Conditions 2, 3, and 4. With the exception of one point for one *O,* sensitivity-scores are improved as the number of stimulus-repetitions is increased.

Discussion. The finding of extensive individual differences is not at all unusual. It would be naïve to expect that all normal *O*s have the same

PRONOUNCEABILITY, ATTENSITY, AND INTERFERENCE IN THE COLOR-WORD TEST

By PAUL BAKAN and BURTON ALPERSON, Michigan State University

In the standard form of the Stroop Test (color-word interference test), the *S*s are shown color names printed in ink of incongruent color, *e.g.* the word 'red' printed in blue ink. *S* is to name the ink colors in a sequence of incongruent color-word stimuli. Under these conditions, the time required to call out the ink colors is longer than the time required to name simple patches of color. The word seems to interfere with the naming of the color.[1]

Klein has gone beyond the usual procedure of using color-words, and has shown that the interference effect can be manipulated by varying the verbal text in which the colors are imbedded.[2] Thus, he found increasing amounts of interference as the nature of the verbal text went from nonsense-syllables, to rare words, to common words, to words that imply color (lemon, grass), to color names different from the ink colors, to color names the same as the ink colors (the standard condition). Klein attributes the interference to the capacity of the word to produce arousal of its motor-components, *i.e.*, the disposition to say the word. He calls his capacity the attention-catching or attensive power of the word. The color-word paradigm can be used as a means of assessing the attention-catching power of verbal material and thus as a technique to study variables influencing the attensive properties of verbal stimuli.

According to Klein's theory, a variable that influences the report of verbal material should also influence the interference in the color-word task. Pronounceability is such a variable. Verbal material which is easy to pronounce is more likely to lead to arousal of motor-components involved in saying the words. The greater attensity of such material should result in greater interference.

In this study, the effect of pronounceability on interference is investigated. The relationship between meaningfulness and attensity is also considred; and an independent tachistoscopic measure of attensity of verbal material is described.[3]

---

* Received for publication October 24, 1966.

[1] The test is described in a paper by J. R. Stroop, Studies of interference in serial verbal reactions, *J. exp. Psychol.*, 18, 1935, 643–662. An extensive review of the literature on this test and its derivatives is found in a paper by A. R. Jensen and W. D. Rohwer, The Stroop color-word test: A review, *Acta Psychologica* 25, 1966, 36–93.

[2] G. S. Klein, Semantic power measured through interference of words with color-naming, this JOURNAL, 77, 1964, 576–588.

[3] E. B. Titchener, The term 'attensity,' this JOURNAL, 35, 1924, 156.

METHOD

Stimuli. The pronounceability scales published by Underwood and Schulz formed the basis for the selection of five groups of four stimuli shown in Table I.[4] The first three groups represent nonsense verbal material in increasing order of pronounceability; Groups 4 and 5 represent meaningful verbal material at two levels of pronounceability. Groups 3 and 4 are matched for pronounceability; Group 3 consists of nonsense material and Group 4 consists of meaningful material.

The stimuli, having been selected on the basis of pronounceability and meaningfulness, were then scaled on an independent measure of attensity. It was assumed that words with greater power to catch attention would be recognized more easily in a tachistoscopic presentation—that the tachistoscopic indicator of attensity would reflect all of the variables contributing to the attensity of a stimulus. Thus, if in the present set of stimuli both pronounceability and meaningfulness contribute to attensity, then the tachistoscopic measure of attensity of Group 4 (words) should be greater than that of Group 3 (nonsense) despite the fact that the groups are of equal pronounceability.

The scaling of tachistoscopic attensity is based on the mean recognitive threshold of 54 *S*s who saw all the stimuli (20) in one of five different orders. Each word was presented in an ascending series of 17 trials ranging from 0.01 to 2.0 sec. The four stimuli within each group were presented in succession and always in the same order. Order of presentation of the stimulus-groups was so rotated for each of five groups of *S*s that each group of stimuli appeared first for each group of *S*s. Presentation of the 20 stimuli was preceded by 4 warm-up words.

TABLE I

STIMULI GROUPED IN TERMS OF PRONOUNCEABILITY

	Group				
	1	2	3	4	5
	FJQ	EKL	DAP	WAS	HAT
	RZQ	XAT	ISH	BOY	CAT
	VGJ	TUW	LAR	ITS	MAN
	XFH	ZOJ	FON	ELK	HUM
Mean pronounceability	8.69	6.51	2.49	2.49	1.71

The words were flashed on a light gray wall in a dimly illuminated room with a Polymetric model V-1459A projection tachistoscope. Following each trial *S* wrote down what he saw on a response sheet. The values in Table I represent the mean recognitive times for each word across all 54 *S*s.

Design. Corresponding to the five stimulus conditions there were five groups of 21 undergraduate *S*s. Each *S* was tested with only one set of verbal stimulus-material on the color-word interference task.

Procedure. Each *S* was tested individually. After entering the experimental room, he was asked to read a card containing 80 color names (red, yellow, blue, green) which were printed in black ink. There were 10 rows on the card, each containing 8 color names. The words were randomly ordered with the restriction that no color name could appear twice in succession on a given row.

---

[4] B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Learning,* 1960, 370–372. Low values on this scale represent easy pronounceability.

The second task of the Ss was to identify a series of 80 patches of color (red, yellow, blue, green). These color patches were randomly ordered, 10 rows on a page and 8 patches in a row, again with the restriction that no color could be presented twice in succession in a given row.

Finally, S identified the colors in which words from one of the five stimulus-groups were printed. The order of ink-colors for this card was identical to the order of colors on the color-patch card. A separate randomization was carried out for the order of words on this card (with the same restriction as above). The same randomization series was used across all stimulus groups.

Instructions to S for all cards placed the primary emphasis on speed, although the Ss were also told that their accuracy would be recorded. They were made aware that they would be timed for each card. Following Klein, they were warned against using various interference reducing devices, *e.g.* sing-song reading, squinting, pointing, etc.

The first line of each card was intended to absorb S's comments about and his initial reaction to the task. The actual timing began with the second line of each card.

TABLE II

MEAN RECOGNITIVE THRESHOLD (ATTENSITY) AND INTERFERENCE-SCORE

Pronounceability	Recognitive threshold (sec.)	Interference
1	.0492	6.00
2	.0414	7.19
3	.0410	9.62
4	.0356	13.14
5	.0342	13.76

RESULTS

The main analysis was based on an interference score, obtained for each S, by subtracting the reading time (sec.) for the non-interference color-patch card from the reading time for the experimental interference card. The mean interference score is shown for each of the stimulus groups in Table II.

An analysis of variance of these interference-scores for the five groups showed differences between the groups significant at the 0.005 level ($F = 12.07$, $df = 4/100$). Individual comparisons (Duncan's multiple range test) reveal significant differences between Groups 1 (low pronounceable nonsense syllables) and 3 (high pronounceable nonsense syllables) and between Groups 3 and 4 (low pronounceable words). Groups 3 and 4 had been matched for pronounceability in order to allow for an evaluation of the effect of meaning on interference.

The mean tachistoscopic recognitive threshold for each of the stimulus-groups is shown in Table II. It is assumed that this gives a more comprehensive measure of the attensity of the stimuli than pronounceability alone. The ordering of the mean recognitive threshold corresponds to the ordering of mean pronounceability with one exception. Whereas Groups 3 and 4 are equal in pronounceability, they are not equal in mean recognitive threshold, the recognitive time being shorter for words than for syllables. The ordering of the mean tachistoscopic thresholds of recognition corresponds perfectly with the ordering of the interference scores for the five groups. This is seen in Table II. The correlation between the pronounceability score and recognitive time for all 20 stimuli is 0.77, significant at the 0.01 level.

An analysis of variance of the between groups differences in tachistoscopic recognitive times shows differences significant at the 0.005 level ($F = 10.12$, $df = 4/265$). Individual comparisons among means yielded significant differences between Group 3 and Groups 4 and 5.

DISCUSSION

Klein has defined the attensive power of a word in terms of its capacity to produce arousal of its motor-components. The attensive power should influence the amount of interference shown in the color-word task. It was hypthesized that a word (or syllable) which is difficult to pronounce should be less likely to arouse its interfering motor-components, and, therefore, it should produce less interference in the color-word task. For three groups of nonsense-syllables (Groups 1, 2, and 3) varying in pronounceability, the ordering of interference scores corresponds with the ordering of pronounceability. For two groups of meaningful words (Groups 4 and 5) varying in pronounceability, the ordering of interference-scores corresponds with the ordering of pronounceability. The hypothesized relationship between pronounceability and interference is confirmed by the experimental results.

Of special interest is the comparison of Groups 3 and 4. By design, these groups of stimuli were equated for pronounceability in order to determine whether meaningfulness would have an effect on interference, with pronounceability controlled for. It was found that the meaningful words of Group 4 produced more interference than the nonsense-syllables of Group 3, despite equal pronounceability. The lack of perfect correspondence in the ordering of pronounceability and interference can be attributed to the contribution of meaningfulness to the attensive value of stimuli as one goes from Group 3 to Group 4.

Klein's assumption that differences in interference scores reflect differences in the attensive value of the stimuli would be supported if a measure of attensity in a context other than the color-word task could be shown to be related to the amount of interference. This experiment points the way to this possibility. All of the stimuli in the experiment were scaled in terms of tachistoscopic recognition. The mean threshold was computed for each of the five sets of stimuli used in the experiment. It was found that the ordering of the groups in terms of interference-scores was in perfect correspondence with the mean threshold of recognition for each group. The tachistoscopic measure of attensity reflects the operation of both pronounceability and meaning, and thus makes for better correspondence of the threshold and interference measures of attensity. The relationship between an independent measure of attensity (recognitive threshold) and the amount of interference offers strong suport to Klein's analysis of the dynamics of the color-word task.

SUMMARY

Verbal material at four levels of pronounceability served as the context for the task of interference. There were five groups of 21 *S*s. Three groups were tested with nonsense-syllables, each with a different level of pronounceability, and two groups were tested with words, each with a different level of pronounceability. Two of the groups worked with material equated for pronounceability, but varying in meaningfulness (words and nonsense-syllables). The *S*s named a series of colored patches, and also named the color of ink in which a series of words (or nonsense-syllables) was printed. The discrepancy between the time taken to name the color patches and the time taken to name the colors in a verbal context was taken as a measure of interference. There was general support for the hypothesis that amount of interference increases with pronounceability. Of the two conditions equated for pronounceability but varying in meaningfulness (words and nonsense-syllables), the interference was greater for the more meaningful material. A tachistoscopic measure of the attention-catching power of the stimuli (attensity) showed a better relationship to interference than the measure of pronounceability. The data are in agreement with Klein's analysis of the dynamics of the color-word task.

THE EFFECT OF LEARNING STRATEGIES UPON FREE RECALL

By Morris N. Eagle,
Yeshiva University

Psychologists carrying out studies in learning and memory have shown a traditional reluctance to ask subjects (*S*s) what they were doing in trying to learn the material presented to them. Instead, they have tended to rely solely on manipulation of factors such as instructional variables and have assumed implicitly that the *S*s do what they are instructed to do. Recently, there has been a tendency to accept the legitimacy and value of obtaining *S*'s verbal reports of how he learned a task as well as a recognition of the ubiquity of *S*'s tendency to employ various learning strategies in trying to learn presented material.[1]

Most recent studies on learning strageties have been concerned with paired-associate learning.[2] In a few studies, the effect of sentential mediation on serial learning was also investigated and was found to be greater than simple naming.[3] There have been relatively few studies, however, in which the relationship between learning strategies and free recall of meaningful words is investigated.

The present paper reports two studies on the effects of rehearsal and associative organization on free recall of meaningful words. Previous work had shown that these two types of strategy were most frequently reported by *S*s.[4] In the first experiment, recall performances of *S*s reporting these two strategies were compared. In the second study, an attempt was made to test the effects on recall of 'instructing' *Ss* to employ either of these two strategies.

---

[1] B. J. Underwood and R. W. Schulz, *Meaningfulness and verbal learning,* 1960, 296–300; G. A. Miller, Eugene Galanter, and K. H. Pribram, *Plans and the Structure of Behavior,* 1969, 125–138; I. E. Farber, The things people say to themselves, *Amer. Psychologist,* 18, 1963, 185–197.

[2] L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, this JOURNAL, 73, 1960, 22–40; W. N. Runquist and F. H. Farley, The use of mediators in the learning of verbal paired associates, *J. verb. Learn. verb Behav.,* 3, 1964, 280–285; W. E. Montague, J. A. Adams, and H. O. Kiess, Forgetting and natural language mediation (NLM), *J. exp. Psychol.,* 72, 1966, 829–833; C. J. Martin, D. L. Cox, and F. J. Boersma, The role of associative strategy in the acquisition of paired-associate material: An alternative approach to meaningfulness, *Psychonom. Sci.,* 3, 1965, 463–464; W. F. Montague and A. J. Wearing, The complexity of natural language mediators and its relation to paired-associate learning, *Psychonom. Sci.,* 7, 1967, 135–6.

[3] A. R. Jensen and W. D. Rohwer, Jr., Verbal mediation in paired-associate and serial learning, *J. verb. Learn. verb. Behav.,* 1, 1963, 346–352.

[4] M. Eagle and E. Leiter, Recall and recognition in intentional and incidental learning, *J. exp. Psychol.,* 66, 1964, 58–63.

EXPERIMENT I

Procedure. A group of 43 women and 9 men students from the Mills College of Education and from Brooklyn College class in introductory psychology served as *Ss* in group testing for the first experiment.

All the *Ss* were presented a list of 20 words taken from the Kent-Rosanoff Word Association Test. They were told simply to try to remember the words as they would later be asked to recall them. The words were read at the rate of one every 3 sec. One minute after the list was read the *Ss* were asked to write their sex on the data sheet, the number identifying them, and as many of the stimulus-words as they could recall. They were given 2 min. for this task. Recall sheets were then collected and the *Ss* were asked to respond briefly to the following questions:

(1) Did you try to remember the words while they were being read to you?

(2) Please describe as specifically as possible any way in which you tried to memorize the words. Please give examples.

Results. The Ss' reports of the methods they used in trying to memorize the words were categorized by *E* and another judge who were both unaware of *Ss'* recall scores. The reports fell primarily into two categories—rehearsal and associative organization. (There were some *Ss* who reported no strategy and also scattered reports of idiosyncratic strategies which could not be categorized as either "rehearsal" or "associative organization." The data for these *Ss* are not included in the results). Agreement between the two judges was perfect.

Of the 52 *Ss*, the 27 reporting a rehearsal-strategy recalled a mean of 10.04 words ($SD = 3.53$), while the mean number of words recalled by the 25 *Ss* reporting an associative organization-strategy was 12.84 ($SD = 2.59$). An analysis of variance for unequal *N*s (Winer)[5] reveals significant differences for strategy-reports ($F\ (1,51) = 20.63$, $p < 0.01$) and no significant sex-differences or interaction-effects.

EXPERIMENT II

Subjects. Group testing was conducted in a large introductory psychology class of 59 women and 39 men at Brooklyn College.

Procedure. The class was randomly divided into three equal groups: (a) no strategy-instructions; (b) rehearsal strategy-instructions; and (c) associative instructions.

The *Ss* were told they were going to participate in an experiment on memory and were asked not to discuss the written instructions given to them. Each *S* received one of the three following sets of instructions:

---

[5] B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 222–227.

Instructions: The *no-strategy* group was simply told that they would be presented a list of words and that they should try to remember these words because they would be asked to recall them afterwards.

The *rehearsal*-group was told to report the words "over and over again to yourself as much as possible."

The *associatively organized* group was told to "associate or organize words that seem to belong together." An example was given and *S* was told that because the words in the list would often not belong together in an obvious way, it would be "very helpful to provide your own associations or organization among the words."

The rest of the procedure followed Experiment I—that is, recall and responses to the questionary.

Results. The *Ss*' reports of strategies they employed were categorized into *rehearsal* and *associative organization* by two judges. There were only a few scattered idiosyncratic reports that could not be categorized as either "rehearsal" or "associative organization." Agreement between judges was perfect except for 3 *Ss*, and the disagreements for these were resolved by discussion.

TABLE I

MEANS AND SDs OF NUMBERS OF WORDS RECALLED BY DIFFERENT STRATEGY-INSTRUCTIONS AND BY DIFFERENT REPORTED STRATEGIES

	Reported strategy								
Instructed strategy	Rehearsal			Associative Organization			Total		
	N	M	SD	N	M	SD	N	M	SD
No Strategy	11	10.46	4.51	22	12.95	2.48	33	12.12	3.46
Rehearsal	15	9.67	3.56	15	12.40	3.46	30	11.03	3.44
Associative Organization	8	9.75	1.57	27	13.07	1.96	35	12.31	2.35
Total	34	9.94	3.80	64	12.88	2.57	98	11.86	2.78

Table I shows the mean number of words recalled in the different instructional groups as well as by strategy reported. (Because there were no differences in recall between men and women, their scores were combined.) An analysis of variance for unequal *N*s (Winer) indicates that strategy instructions *per se* have no significant effect on the number of words recalled ($F\ (2{,}92) = 0.34$). Instructions do, however, influence the *number* of *S*s reporting the strategy they are instructed to employ. It can be noted in Table I that of the 30 *S*s instructed to rehearse, 15 reported rehearsal-strategy and 15 reported associative organization. Of the 35 *S*s instructed to employ associative organization, only 8 reported rehearsal and 27 reported associative organization. The differences between these two groups are statistically significant ($X^2 = 4.09$, $p < 0.05$).

It is the strategy employed or at least reported, however (as indicated in *Ss'* verbal reports), that is the critical variable in influencing amount recalled (F (1,92) = 18.03, $p < 0.01$). Furthermore, there are no interaction-effects. The mean number of words recalled for the different reported strategies is strikingly similar across instructional groups. In all instructional groups, the use of associative organization leads to greater recall than does rehearsal.

Discussion. It is clear from the results that strategy-instructions *per se* have no *direct* effect on recall. What instructions accomplish is to shift the number of *Ss* reporting strategies in accord with the instructions. It is the strategy that *S* actually employs, however,—whether instructed to or not—that is the critical factor in influencing amount of recall. This is not surprising if one considers that it is not what *S* is told to do, but what he actually does, that is the critical variable in learning. Even if instructions in strategy were significantly related to recall, it would not be attributable to the instructional variable *per se,* but to the fact that a sufficient number of *Ss* used the strategy they were instructed to use.

Since the results reported here are based on a correlation performance in recall and *Ss'* responses (re: strategy) to post-experimental inquiry, it could be argued that recall performance determined strategy reports (rather than, as argued in this paper, that strategy employed determined recall performance). This possibility cannot be ruled out, but seems somewhat implausible. It suggests that after recall, *S* selects a strategy to report which matches his performance. Furthermore, there is evidence in the literature both for paired-associate learning and free recall indicating that items for which *S* can report the use of a strategy (mediator) are retained better at a *subsequent* time.[6] Finally, it has already been shown in this study that strategy-reports are significantly influenced by strategy-instructions. If reports were simply an artifact of performance, one would not expect them to be affected by instructions. Implausible or not, however, work is in progress which will more effectively evaluate this possibility. For example, information regarding *S*'s preferred strategy will be obtained independently of and prior to recall.

As for differential effects of reported strategies, associative organization results in significantly better recall than does straight rehearsal. This finding agrees with Tulving's work on "subjective organization" and also makes theoretical sense.[7]

---

[6] H. O. Kiess and W. E. Montague, Natural language mediators in paired-associate learning, *Psychonom. Sci.*, 3, 1965, 549–550; Montague and Wearing, *op. cit.;* L. D. Groninger, Natural language mediation and correct rehearsal in short term memory, *Psychonom, Sci.*, 5, 1966, 135–136; F. J. Boersma, R. C. Conklin, and J. E. Carlson, Effects of reporting associative strategies on the retention of paired-associates, *Psychonom, Sci.*, 5, 1966, 463–4.

In the strategy of associative organization, a structure is imposed on the stimulus-words whereby the recall of any given word is more likely to cue the next word, with the result that amount of recall is increased. The *S*s' reports indicate that although the most frequent means of creating a structure is the associative linking (two or more words that "seem to belong together" are associated together), the *S*s do report other techniques —*i.e.,* syntactical linking (words formed into a phrase or sentence), or some other kind of combination (*e.g.* words combined to form an imagined scene). What these techniques appear to have in common is that *S*s are, in effect, "chunking" words to a greater degree than *S*s employing rehearsal.[8] Two or more words are organized into larger units with the result that the informational content in each unit is enriched. If, as Miller indicates, memory-capacity is limited by the number of units rather than the amount of information in each unit, one should expect better memory performance to result from the "chunking" of words into associative, syntactical or other units.

SUMMARY

Two studies were carried out on the relationship between learning strategies and recall. In the first study, *S*s who reported using an associative organization, *i.e.* strategy—linking words in some manner—recalled significantly more words than *S*s reporting straight rehearsal. In the second study, strategies were assigned via instructions to different groups. Strategy instructions *per se* had no direct effect on recall, but *reported* strategy significantly influenced amount of recall. In all three instruction-groups, *reported* associative organization was associated with greater recall than *reported* rehearsal. The discussion dealt with the concept of "chunking" and the importance of *S*'s reports of what he is doing, in predicting recall.

---

[7] Endel Tulving, Subject organization in free recall of "unrelated" words, *Psychol. Rev.,* 69, 1962, 344–354.

[8] G. A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information, *Psychol. Rev.,* 63, 1956, 81–97.

TEMPORAL ARRANGEMENT AND SPECIFICITY OF SET IN AN AUTOKINETIC SITUATION

By T. E. Dielman and Kenneth H. David, University of Hawaii

The problem investigated by this study was whether a suggestion of stimulus-movement would be more effective when presented before or after the onset of a physical stimulus (temporal arrangement); and whether, given a highly ambiguous stimulus, a more specific response category would lead to a lower response frequency (specificity of set).

Concerning the temporal arrangement of set, Garner, Hake, and Eriksen stated that if the facilitative effect of set is due to modification of perception, then the effect should occur only when alternatives are given prior to the stimulus and should not occur when given after exposure to the stimulus.[1] Lawrence and Coles, using tachistoscopically presented familiar objects, found no differences between groups of *Ss* presented with prior set or subsequent set.[2] Long, Reid, and Henneman, using degraded letter-stimuli, found that a set presented prior to the onset of the physical stimulus had a significant effect only when the *Ss* were given two alternatives from which to choose.[3]

In regard to specificity of set, Long, Reid, and Henneman found that fewer alternatives resulted in a greater number of responses.[4] Lawrence and Coles found that when the alternatives were discrete (a less ambiguous stimulus-situation), the *Ss* emitted more responses than when the alternatives were similar (a more ambiguous stimulus-situation).[5]

In the present study, a variation of the autokinetic word-technique devised by Rechtschaffen and Mednick was used.[6] The *Ss* were seated in a darkened room and told that a stationary pinpoint of light was going to move in a certain manner. Temporal arrangement of set was defined as the suggestion of stimulus-movement either before or after the onset of the stimulus-light. Specificity of set was defined by the instructions to the *Ss* that the light would trace either vowels (high specificity or few alternatives) or conso-

---

* Received for publication August 8, 1966.

[1] W. R. Garner, H. W. Hake, and Charles W. Eriksen, Operationism and the concept of perception, *Psychol. Rev.*, 63, 1956, 149–159.

[2] D. H. Lawrence and G. R. Coles, Accuracy of recognition with alternatives before and after the stimulus, *J. exp. Psychol.*, 47, 1954, 208–214.

[3] E. R. Long, L. S. Reid, and R. H. Henneman, An experimental analysis of set: Variables influencing the identification of ambiguous, visual stimulus-objects, This Journal, 73, 1960, 553–562.

[4] *Ibid.*

[5] Lawrence and Coles, *op. cit.*

[6] Allen Rechtschaffen and S. A. Mednick, The autokinetic word technique, *J. abnorm. soc. Psychol.*, 51, 1955, 346.

nants (low specificity or many alternatives). Control groups were included, in which the *S*s were merely told that the light would move. The experimental design was a 2 x 2 x 3 (Sex × Temporal Arrangement × Specificity of Set) factorial design.

It was hypothesized (1) that the suggestion of a stimulus-movement before the onset of the stimulus-light would result in a greater number of responses than a subsequent suggestion of stimulus-movement; and (2) that in the highly ambiguous autokinetic situation, the more specific response category (vowels) would result in fewer responses than the less specific response category (consonants).

Method: Subjects. The *S*s were 120 undergraduate students (60 men and 60 women) from Introductory in Psychology at the University of Hawaii.

Apparatus. The apparatus consisted of a small, frosted, 5-w. bulb enclosed in a fiberboard box, the only source of light being a 1/32 in. hole in the front of the box, with masking tape placed over the hole. The light and a variable-pitch motor (color-wheel) were operated by a common switch. Black cloth was used to cut off seepage of light around the box and under the door. The *S*s were seated 6 ft. from the source. The visual angle subtended was 0.008°. The intensity of the light as determined by a monocular brightness match procedure was −1.20 log ML.

Procedure. The *S*s were assigned to the six conditions on a random basis, with 20 *S*s (10 men and 10 women) to each condition. They were met by *E* in a waiting room and told that the instructions would be given once they were seated in the experimental room. They were then taken to the experimental room (6′ × 12′ with black interior) and seated by *E*. The instructions were then given them, after which *E* switched on the stimulus-light and motor and turned on a timer that was pre-set for 3 min. At the sound of the timer-alarm, the light was switched off and the *S* presented the post-stimulus instructions.

Instructions: The instructions given the *S*s were as follow:

Prior-set: This is an experiment designed to test tracking ability under various conditions of dark-adaptation. I have a variable pitch projector with which I am able to trace letter patterns on the wall before you. I will turn on the projector and give you a 3 min. trial during which I will be writing vowels (consonants). Simply count the number which you see and wait till after the test to report them to me. [After the three-minute trial the *S*s in the prior set condition were asked to report the number of vowels (consonents) they had seen written.]
Post-set: [The *S*s in this group were seated in the darkened room and told]: This experiment is designed to test tracking ability under various conditions of dark-adaptation. I have a small, variable pitch projector with which I will be projecting a moving light on the wall before you. Your task is to watch the light as closely as possible during 3 min. and report to me after the trial what you have seen. [Upon completion of the trial they were told]: Without telling you beforehand what I was going to do, I wrote vowels (consonants) with the moving light. How many did you see?

Control group. The *S*s in the control group were given the pre-trial, post-set instructions. After 3 min. they were asked to "report what you have seen."

After *S* reported the number of vowels (consonants) he had seen or what the light had appeared to do, he was taken back to the waiting room to complete the "awareness" questionary (*e.g.* "What do you think the purpose of this experiment was?") and asked not to discuss the experiment with anyone until the end of the semester.

Care was excercised throughout the experiment to insure that the *Ss* could not see anything in the experimental room either before or after doing the task.

Results and discussion. The mean number of responses made by each group during the three-minute session is presented in Table I. The Prior-set consonant *Ss* had the highest mean score (3.55) and the Prior-set vowel *Ss* had the second highest mean score (2.45). The other four groups obtained mean scores that were 0.75 or less. Two *Ss*, both in the Prior-set group, indicated "awareness" on the questionary. Neither of these *Ss* reported seeing consonants and their scores were kept in the analysis.

The results of the analysis of variance indicated that the presentation of suggestion of movement prior to the onset of the stimulus resulted in a significantly greater number of responses than the suggestion after the stimulus ($p < 0.01$). The specificity factor showed a significant difference ($p < 0.05$), and the Newman-Keuls procedure was used to determine where the differences occurred. The Newman-Keuls procedure indicated that both the Consonant group and Vowel group were significantly different from the Control group ($p < 0.05$); however, there was not a significant difference between the Consonant and Vowel groups. There were no sex differences or interaction effects.

TABLE I

MEAN NUMBER OF RESPONSES

($N = 20$ per cell)

Temporal arrangement	Specificity of set		
	vowel	consonant	control
Before	2.45	3.55	.40
After	.75	.50	.00

The results supported the hypothesis that the suggestion of stimulus movement before the onset of the stimulus light would result in a greater number of responses than a subsequent suggestion. This supports the position that perception is modified by suggestion and that *Ss* do not merely modify their responses to conform to the suggestion of *E*. In other words, the expectation of stimulus-movement appears to have altered *Ss*' perception of the stimulus.

The results failed to support the second hypothesis, that the specificity of the response category would be related to the number of responses.

Summary. The temporal arrangement of set and the specificity of a response-category was varied in a three-way design (Sex × Temporal Arrangement × Specificity) using 120 *Ss* (60 men and 60 women). A variation of the autokinetic word-technique was used. Temporal arrangement of set was defined as the suggestion of stimulus-movement either before or after the onset of the stimulus-light. Specificity of set was defined by the instructions that the light would trace either vowels (high specificity or few alternatives) or consonants (low specificity or many alternatives). A control condition was included, in which the *Ss* were merely told that the light would move. The results indicated that the presentation of suggestion of movement prior to the onset of the stimulus resulted in a significantly greater number of responses than the suggestion after the stimulus ($p < 0.01$), which supports the position that perception is modified by suggestion. The results failed, however, to support the hypothesis that the specificity of the response-category would be related to the number of responses.

RETRIEVAL TIME AND THE EFFECT OF SERIAL POSITION

By ALAN KENNEDY and ANNE GALLOWAY, Queen's College,
Dundee, Scotland.

The theory of learning put forward by Feigenbaum sees the effect of serial position as arising directly from the strategy of encoding and retrieval used by the learner.[1] This theory provides a good prediction of the general form of the serial position-curve as described by McCrary and Hunter.[2] Although the position taken by McCrary and Hunter probably oversimplifies the operation of those factors leading to the effect of serial position,[3] Feigenbaum's theory at least provides a number of testable predictions. The model assumes serial lists to be processed beginning with perceptually unique anchor-points: these normally would be the first and last items of a list. During learning, a branching discrimination-net is built up which comprises a series of choice-points providing unique pathways to a number of terminal nodes at which are stored coded stimuli and also limited information or cues to enable related responses to be retrieved. The relative difficulty of items near the middle of a serial list can be seen to follow from the fact that they are processed later in learning and are retrieved after having been sorted through a relatively greater portion of the discrimination-net. During learning, this gives rise to the slower acquisition of these items, with a higher probability of errors or omissions. For the same reasons, during recall of a completely learned list, items in the center should show a longer response-latency than those at either end. This last prediction has not been previously tested, and the present study sets out to examine it, using two simple learning tasks.

The learning situation was similar to that used by Murdock.[4] A list of paired words was presented and after an interval *S* was given a single stimulus-item and asked to provide the response. Only one presentation was used, but *S* was given 1 min. in which to organize and rehearse the list. To avoid the effects of proactive inhibition found when the same *S* is used for several different lists, only one list was presented to each *S*.[5] The situation was considered to be as free from complicating factors as possible: a single presentation to each *S*; no attempt to inhibit rehearsal; and one position only sampled within each list.

---

[1] E. A. Feigenbaum, An information processing theory of verbal learning, 1959, RAND *Report P-1817*, Calif.; The simulation of verbal learning behaviour, *Proc. West. Joint Computer Conf.*, 1961, 19, 121–132; E. A. Feigenbaum, and H. A. Simon, A theory of the serial position effect, *Brit. J. Psychol.*, 53, 1962, 307–320.

[2] J. W. McCrary and W. S. Hunter, Serial position curves in verbal learning. *Science*, 117, 1953, 131–134.

[3] A. R. Jensen, Is the serial position curve invariant? *Brit. J. Psychol.*, 53, 1962, 159–166.

[4] B. B. Murdock, Short-term retention of single paired-associate items, *Psychol. Rep.*, 8, 1961, 280.

Two lengths of list were used in the present study to investigate the further prediction that, regardless of serial position, latencies should be longer over-all for a longer list.

Method: (1) Design. The original conception of the experimental design was as a 3 × 3 × 2 factorial, with serial position, order of presentation, and list-length as relevant variables. While two separate word-lists were used, it had been first intended to employ the same key pairs in both lists; in this way the three-pair list was to be converted into a five-pair list by the addition of two pairs in the second and fourth position. The three positions sampled would have then involved the recall of the same words for both list lengths, but since a five-pair list was rarely learned in one trial, a very high rejection rate of *Ss* made the procedure suspect. Thus it was not possible to incorporate list-length into an over-all design and the experiment was analyzed in two separate parts. The three positions sampled were the first, middle, and last, and the lists were presented in three random orders so that each word pair occurred at a different position.

Materials. Two lists were used: a three-pair list from the word-lists of Greenberg and Underwood (1950),[6] and a five-pair list of easy related pairs. The lists were: (a) Joyful–Gallant; Brittle–Pious; Latent–Drastic; (b) Foolish–Silly; Tranquil–Peaceful; Awkward–Clumsy; Barren–Sterile; Vacant–Empty.

Lists were recorded on a standard Ferrograph tape recorder. A voice-key using a sensitive trigger and appropriate steering circuitry was so used that stimulus-items started a timer and *S*'s spoken response stopped it.

Subjects. The *Ss* were 75 women, students in college and university training. None was acquainted with the techniques used in the experiment. The *Ss* were assigned to experimental conditions at random with the restriction that no cell would be wholly filled during one testing session. After discarding those who failed to respond or gave an incorrect response, the data of 54 *Ss* were available for analysis.

Procedure. A color-naming warm-up task preceded the test for every *S*. The following instructions were then read:

> You will hear a list of pairs of words once on the tape recorder. Please learn the pairs so that when you hear the first word of a pair you can give the second word which completes that pair. At the end of the list there will be an interval of about 1 min., then you will be told 'ready'. After a short pause you will hear the first word of one of the pairs. Will you please say the second word of that pair as quickly as you can, speaking directly into the microphone in front of you. Remember to reply as quickly as you can and speak clearly. Do not say anything before you say the second word.

Lists were presented at the rate of one word per sec. with 2 sec. between successive pairs. The interval between presentation and recall was 1 min., and the warning 'ready' was given 5 sec. before the cue word for recall. No attempt was made to prevent rehearsal between presentation and recall.

---

[5] G. Keppel and B. J. Underwood, Proactive inhibition in short-term retention of single items, *J. verb. Learng. and verb. Behav.*, 1, 1962, 153–161.

[6] R. D. Greenberg and B. J. Underwood; Retention as a function of stage of practice, *J. exp. Psychol.*, 40, 1950, 452–457.

Results. (1) Lost S. Since an inevitable feature of the experimental situation was a high wastage of *S*s, a careful analysis of their distribution in the experimental conditions is called for. Of 21 *S*s discarded, 4 were lost through apparatus-failure and 3 others misunderstood the task or failed to give a response. Of the 14 incorrect responses given there was no tendency for these to fall in any one position: for the three positions, losses were in the three-pair list 2, 4, and 4; and for the five-pair list, 1, 1, and 2.
(2) Three-pair list. Mean latencies ($N = 9$) for the three positions were 1.55, 1.74, and 1.17 sec. The results were analyzed with order of presentation and position as factors. Position was significant ($F = 5.03$; $df = 2, 18$; $p < 0.025$). Positions 1 and 3 ($p < 0.05$) and 2 and 3 ($p < 0.01$) both differed significantly. For this list, order was also marginally significant. ($F = 3.54$ where $F = 3.55$ with $df = 2, 18$, is significant $p < 0.05$) there being a slight tendency for latencies in one order of presentation to be longer than average. There was no apparent reason for this on inspection and the effect must be put down to some special property of the combinations of the pairs. Order and position did not interact, ($F = 1.79$, $df = 4, 18$).
(3) Five-pair list. Mean latencies ($N = 9$) for the three positions were 1.49, 1.70, and 1.20 sec. There is clearly no tendency for latencies in this list to be longer in general than those of the three-pair list. Order of presentation was not shown to be significant ($F = 0.005$, $df = 2, 8$), but position was significant ($F = 4.20$, $df = 2, 18$, $p < 0.05$): positions 2 and 3 differed significantly ($p < 0.01$). There was no tendency for position and order to interact.

Discussion. The failure to find a longer latency for the longer list may be explained by considering the composition of the five-pair list. It has been shown in an object-naming task that response latency for high-frequency words is shorter than for low-frequency words.[7] The words of the five-pair list occur slightly more frequently than those of the three-pair list and when it is considered that the response is already known to the *S* as part of the associative cluster around the stimulus-word, it is possible that shorter latencies could be found for items of this kind. Any effect in this direction would offset the tendency for latencies to be longer for the longer list.
The results of the analysis of position-effects for both lists are in broad

---

[7] R. C. Oldfield and A. Wingfield, Response latencies in naming objects. *Quart. J. exp. Psychol.*, 17, 1965, 273–281.

agreement with Feigenbaum's theory of verbal learning in that consistent time differences occur in retrieving different parts of sequentially presented material. This is a "serial position effect" of a rather different type from that shown using error-scores, but it illustrates the possible basis for the latter effect. While detailed predictions from Feigenbaum's theory would here lead to the expectation of shorter latencies for the first items in a list, the general picture is of a storage system which much be searched from either end. Analysis of the serial position effect has usually remained at the descriptive level that items near the center are "more difficult." The present results give some meaning to this description when applied to a well learned, easy, list. It appears on the basis of this measure that items near the center of a list are more deeply embedded in the storage system, and consequently are less easily retrieved.

REDUCTION OF AUTOKINESIS WITH INFORMATION ABOUT THE REGISTRATION OF EYE-POSITION

By STEPHEN WORCHEL and CLARKE A. BURNHAM, University of Texas

Recent investigators have proposed that the autokinetic effect results from a misinterpretation of the record of the efferent commands sent to the extraocular muscles to maintain fixation on the light.[1] This interpretation is supported by the finding that when the eye muscles are strained, autokinesis occurs predominantly in the same or opposite direction as the previously induced strain; such movement reflects the abnormal efference required to maintain fixation in these directions.[2] As Royce, Carran, Aftanas, Lehman, and Blumenthal point out, this explanation does not, however, account for the effect of suggestion on autokinesis or why autokinesis diminishes with the establishment of a frame of reference and is not seen in normal daylight viewing.[3]

We suggest that there are several available sources of information which are capable of providing input regarding the efficiency of the monitoring of the efferent signals and hence capable of altering the amount of their misinterpretation by the nervous system. The presence of assumed stable objects may prevent this misinterpretation in normal daylight viewing. Suggestions about the magnitude of autokinesis may facilitate misinterpretation by causing *S* to think that the monitoring is inefficient. The present study represents an attempt to reduce the magnitude of autokinesis by providing *Ss* with a non-visual and non-verbal source of information about the possible range of their eye movements. It was hypothesized that if the autokinetic light was viewed through a tube which was felt by *S* but which remained non-visible, the time to onset of movement would be longer and the magnitude of movement would be less than in an unrestricted viewing condition. The presence of the tube should serve as a reference input and

---

* Received for publication January 26, 1967. This research was supported in part by a grant from the NSF Undergraduate Research Participation Program, GY-1021.

[1] R. L. Gregory and O. L. Zangwill, The origin of the autokinetic effect. *Quart. J. exp. Psychol.*, 15, 1963, 252-261; Irvin Rock, *The Nature of Perceptual Adaptation*, 1966, 109.

[2] R. L. Gregory and O. L. Zangwill, The origin of the autokinetic effect, *Quart. J. exp. Psychol.*, 15, 1963, 252-261.

[3] J. R. Royce, A. B. Carran, M. Aftanas, R. S. Lehman, and A Blumenthal, The autokinetic phenomenon: A critical review, *Psychol. Bull.*, 65, 1966, 243-260.

inhibit the misinterpretation of the relationship between the efferent signal and the visual signal.

Method. The experiment was conducted in a light-proof room. The light stimulus was provided by a standard 3-v. flashlight bulb to which 1.2 v. was applied. It was placed behind a ½-in. thick plywood panel and viewed through a 1/32-in. hole. Two viewing conditions were used. In the unrestricted viewing condition, the *Ss* viewed the light monocularly from a distance of 15½ in., with their foreheads and chins supported by a frame. In the restricted viewing condition, a tube was placed between the eye and the light source. The tube had an outside diameter of 11/16 in. and was so positioned that it was flush with the wooden panel and extended to within ½ in. of the *S*'s eye. The *Ss* were unable to see the tube and its interior did not reflect any light. While viewing the light *S*'s right arm was extended and his hand placed directly beneath the light source. In the restricted condition, his thumb and fingers were placed in a U-shaped fashion around the tube. The same positioning of the arm and hand was used in the unrestricted viewing condition. The *Ss* served for three 3-min. trials in both viewing conditions in either a ABABAB or BABABA order. There was a 1-min. interval between trials. The *Ss* were partially dark-adapted during the 5-min. period necessary to explain the instructions. There was no illumination between trials.

The *Ss* were asked to state when the light started to move and to estimate its direction and distance of movement. These verbal reports were tape-recorded for analysis. All the *Ss* were volunteer university students and were totally naive about the autokinetic effect. Six men and six women served as *Ss*.

Results. Two measures of the autokinetic effect were derived from the tape-recorded reports. One was the average time of onset of the initial movement for the three trials in each condition. The other was the average amount of movement for the trials in each condition. The data were analyzed by a 2 × 2 analysis of variance design with correlated scores for the viewing condition. The other factor was the viewing order, unrestricted vs. restricted first. The time to onset of movement was significantly longer in the restricted viewing condition ($F = 6.76$, $df = 1/10$, $p < 0.05$). The mean latency was 25.75 sec. in the restricted condition and 19.37 sec. in the unrestricted condition. The average amount of movement with restricted viewing was 2.40 in.; with unrestricted viewing it was 3.67 in. These figures also differ significantly ($F = 16.22$, $df = 1/10$, $p < 0.01$). There were no significant effects for viewing order or the interaction of order and viewing condition with either measure.

Since the time to onset in the restricted viewing condition was longer, there was less time for movement to occur. Although the difference in onset time between the conditions is relatively small, 6.38 sec., compared to the viewing time of 180 sec., this difference could conceivably account for some of the difference in the total amount of movement. Consequently a third score was derived for each *S*; namely, the amount of movement per time available from the onset of movement. The difference between the conditions with this measure was also significant ($F = 18.15$, $df = 1/10$, $p < 0.01$) indicating that the presence of the tube reduced the amount of autokinesis from the time movement was first perceived as well as increasing the time to onset of movement and reducing the total amount of movement.

A correlational analysis resulted in a substantial positive correlation between the average amount of movement for the two viewing conditions ($r = 0.89$, $df = 10$, $p < 0.01$). The correlation between the average time of onset for the viewing conditions was non-significant ($r = 0.32$).

Discussion. The presence of a non-visible tube through which the stimulus-light was viewed resulted in a significant reduction of the autokinetic phenomenon. It is unlikely that the pattern of results is due to a direct influence of the presence of the tube on the *Ss'* verbal reports of movement. On all but one trial for one *S* the reported pattern of movement would have taken the light beyond the outside diameter of the tube. The high correlation between the amount of movement perceived in the two conditions implies a general reduction of autokinesis for all *Ss* rather than a selective reduction which would presumably result from an influence on reported rather than perceived movement.

The relationship between eye movements and autokinesis is ambiguous.[4] The presence of the tube in this study would not have prevented or hindered eye-movements; their frequency and magnitude would be expected to be similar in both experimental conditions. If eye-movements do play a determining role in the autokinetic phenomenon, the appearance of autokinesis must still result from a lack of accurate registration of the efferent signals to the extraocular muscles, since the position-sense of the eye is determined by outflow rather than inflow information.[5] Thus the information provided by the tube in the restricted viewing condition of this experiment would function in a similar manner irrespective of the role of eye movement in autokinesis. Our experimental results show that providing information about the possible range of the position of the eyes reduces autokinesis; we have hypothesized that this occurs because the tube inhibits a decalibration or misinterpretation of the relationship between the registered position of the eyes and the retinal location of the light.

---

[4] Royce, Carran, Aftanas, Lehman, and Blumenthal, *op. cit.*, 243–260.

[5] G. S. Brindley and P. A. Merton, The absence of position sense in the human eye, *J. Physiol.*, 153, 1960, 127–130; Leon Festinger and L. K. Canon, Information about spatial location based on knowledge about efference, *Psychol. Rev.*, 72, 1965, 373–384.

SUMMARY

The autokinetic effect was significantly reduced when subjects viewed the stimulus light through a non-visible tube. The results were interpreted as supporting the hypothesis that the autokinetic effect results from a misinterpretation of the relationship between the efferent signals to the extraocular muscles and the retinal location of the light. The tube provided information about this relationship and inhibited its misinterpretation.

APPARATUS

AN APPARATUS FOR DETECTING AND MONITORING MOVEMENT

By Edward Griffiths, Norman Chapman and Dugal Campbell,
Queen's University, Kingston

This note describes the design and some functions of an apparatus which will detect and monitor movement.[1] As a detective device, it can report that a movement has occurred (in this respect resembling a stabilimeter), or that some body has entered or left a designated area (in this respect resembling a microswitch or photoelectric cell). As a monitoring device, it can report on the extent and rate of movement (in this respect resembling a strain-gauge or mechanical transducer). The apparatus is small, robust and cheap; and it is not necessary to attach it to the object or person being monitored.

A variety of models can be made; the characteristics of each model vary due to a number of factors described below. For any given application, some trial and error is necessary to arrive at the best combination of parts. The essentials of the apparatus are most conveniently built as two units: a control box and a movement sensing unit. The control box contains a power supply, a variable oscillator, a rectifier diode and filter capacitor, and the necessary input and output connections. The unit registering movement is joined to the control box by cables which serve two flat-wound coils which are, respectively, the input coil and the output coil.

The principle of operation is illustrated in Fig. 1 showing the use of the apparatus as a stabilimeter for infants. The oscillator drives a high-frequency current in the input coil. The space or the object to be monitored and the output coil both lie within the field set up by the oscillating current. A movement in the field causes a momentary change in the voltage induced in the output coil. These alterations in voltage constitute the output of the movement sensing unit.

A considerable variety of arrangements is possible in constructing the apparatus which can be adapted to the requirements of the experiments in view. Changes in the size and winding of the coils, the size and separation of the two coils, and the frequency of the current fed to the coil are among the factors to be taken into account. For example, a field extending about 2 ft. from the input coil, with a maximal width of 8–10 in., is developed by two circular coils placed about 1 in. apart when the input coil is about 3 in. in diameter wound with an interval of 1/16th in. between turns and a similar output coil is about 2 in. in diameter. Smaller coils, 1 in.

---

[1] The development of the apparatus described in this note was supported by the Ontario Mental Health Foundation (Grant No. 16).

across, wound in the same way and also put 1 in. apart, develop a field about 6 in. deep and 2–3 in. wide. The frequency of the current fed to the input coil influences the sensitivity of the device; maximal sensitivity is most readily obtained by tuning the oscillator in conjunction with the type of object it is intended to monitor.

One model of the movement sensing unit is shown in Fig. 2. It was designed to make observations from babies and has been used as a stabilimeter, to provide a count of activity, and as a respirometer, to measure breathing. This model consists of a power-supply delivering 12-v. DC to a solid-state variable oscillator designed to put out frequencies in the range 3–5 MHz. The oscillating current is led to a flat

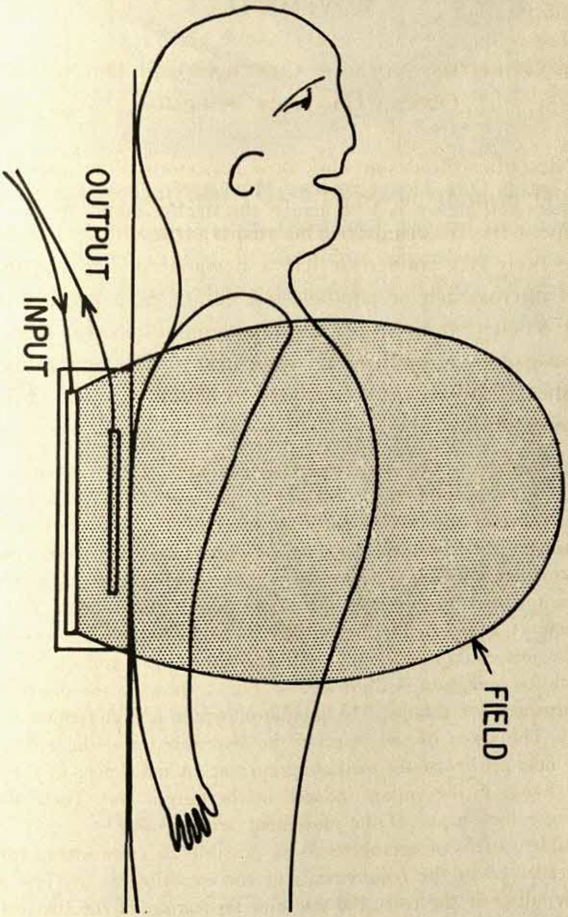


FIG. 1. THE FIELD SET UP BY THE OSCILLATING CURRENT IN THE INPUT COIL IS SHOWN BY THE STIPPLED AREA; THE BODY MONITORED—IN THIS CASE A BABY—AND THE OUTPUT COIL BOTH LIE WITHIN THE FIELD.

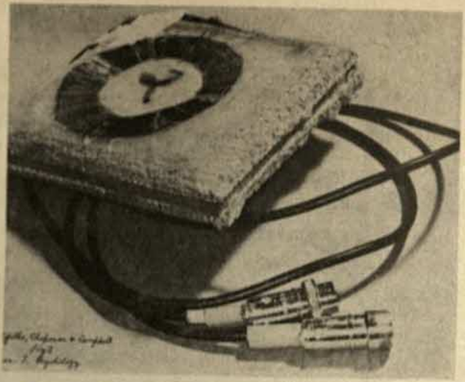


FIG. 2. ONE MODEL OF THE MOVEMENT SENSING UNIT
The larger input coil shown is 3 in. across; the smaller output unit is 2 in. across. The packing material is carpet.

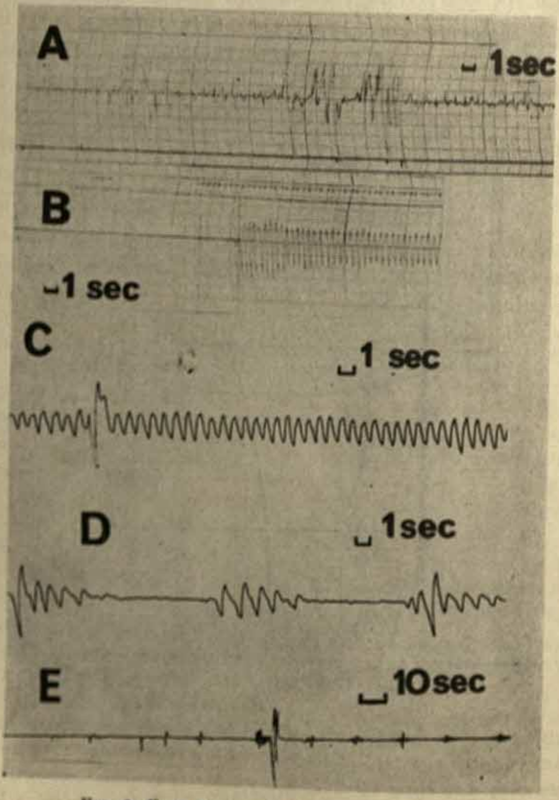


FIG. 3. FIVE RECORDS MADE FROM BABIES
The upper two traces show activity and the end of activity in an infant beginning to suck on a nipple. The lower three tracings show breathing in a premature baby.

input coil about 3 in. across made of fine copper wire coated with insulating varnish and wound with an interval of about 1/16th in. between turns. Each coil is placed in a plastic envelope and the windings are held in place by plastic strips woven through the coil from the center to the edge. The input coil lies in the bottom of a plastic box. The output coil, about 2 in. across, lies just below the top of the box and is separated from the input coil by 1 in. of packing material. The leads from the output coil go directly to a high-frequency diode rectifier which rectifies the voltage developed in the output coil. A filter capacitor is used to filter out the 3–5 MHz. Thus, the resulting output is a direct current potential. In the unit described here the steady potential is 10 v. and movements in the field can cause changes of up to 5 v. These signals may be used to actuate a polygraph or, with suitable amplification, to actuate a counter.

Some specimen results, taken from measures made upon newly-born infants, are shown in Fig. 3. The movement sensing unit used to make these records is shown in Fig. 2; the input current was 12 v. DC at 2.2 M~. The signal has been fed to a writing unit and the five tracings show activity in a normal baby: in this instance the baby was very active (Trace A) and then became quiet when given a nipple (Trace B). The three lower tracings show respiration and activity in a premature baby of 1950 grm. In the first tracing (Trace C) respiration is regular and shows one body movement. Within half a minute, this pattern was replaced by an irregular pattern reminiscent of Cheyne-Stokes respiration (Trace D). The lowest tracing (Trace E) was made with less amplification and shows the change from the regular to the irregular pattern of breathing: first, regular breathing without movement, then the beginning of irregular respiration, a movement and last, more irregular breathing.

A RECORDER OF LOCOMOTOR ACTIVITY

By Roy W. Pickens and William F. Crowder, University of Mississippi

The apparatus described here was designed to register only locomotor behavior and also to minimize sensory feedback from *S*'s activity. It is inexpensive, simple to construct, and easy to service. The activity cage, shown in Fig. 1, consists of a circular alley with a lamp at its center and four photocells spaced equally around its outer wall. To construct a cage, first make a square wooden frame of 1 in. thick boards, 15 × 15 in. inside dimension and about 6-in. high, with a floor of ½-in. hardware-cloth. Then form the circular outer wall of the alley from a strip of sheet-metal, attached to the wooden frame by eight small wood screws. Midway between the ends of each board, drill a ¼-in. hole through the board and the sheet-metal. The holes may center about 1½ in. above the floor for rats or ¾ in. if mice as well as rats are to be used. Paint the sheet-metal and the inside of the holes flat back. Next, make the 9-in. diameter inner wall from a strip of ½-in. hardware-cloth and fasten it to the cage floor

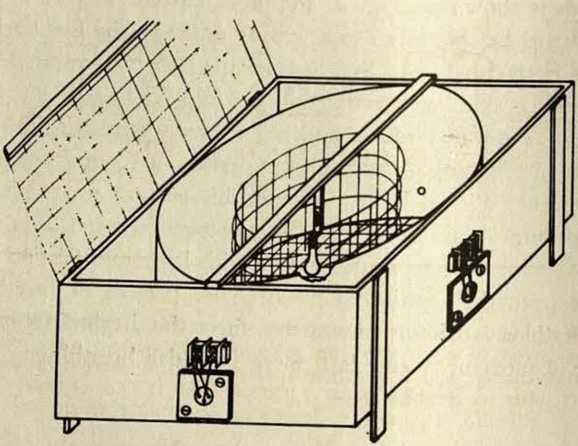


Fig. 1. Circular Activity Cage
One of the two cage lids and two of the four photocells are shown.

---

[*] The development of this device was supported by a University of Mississippi Faculty Research Grant and by Grant No. HD 00916 from the National Institutes of Health.

with short pieces of wire. (The inner wall is made of hardware-cloth instead of sheet metal that the light will illuminate all parts of the alley, rather than only the locations near the photocells. If opaque walls having infra-red filters were substituted, the radiant energy falling on the photocells would be so weak that amplification would be needed in the relay circuit.) A candelabra-base socket for the 6-w., 120-v. lamp is attached by means of a piece of sheet-metal to the wooden strip shown fastened across the top of the cage. The lamp filament should be in line with both opposite pairs of holes. The photocells (Clairex #CL703L/2; Allied Electronics, 100 N. Western Ave., Chicago, Ill.; $1.80 each) are dual-element cells and should be wired with elements in parallel, *i.e.* with the two outside leads connected together. This arrangement gives the cell the lowest resistance when illuminated of any small, inexpensive photocell known to the writers. The photocells mount in 1 × 1 × ¼-in. thick blocks, cut from aluminum bars. Each block should be drilled with a 9/32-in. drill and then carefully reamed with a 5/16-in. adjustable reamer to provide a tight press fit for its photocell.

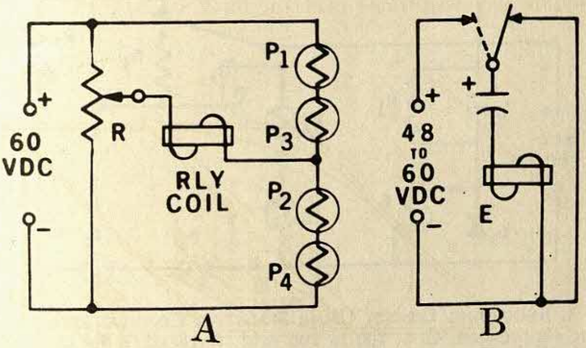


FIG. 2. RECORDING CIRCUITS USING POLARIZED MAGNETIC LATCHING RELAY (A) Bridge circuit of photocells and potentiometer operates relay coil. (B) Relay contact circuit; counter coil E is actuated when capacitor charges or discharges, each time contact shifts its position. R = 2,000 to 5,000 Ω (not critical). C = 100 to 200 mfd. (typically), depending on counter sensitivity.

Fig. 2 shows the recording circuits. The photocells and potentiometer together form a bridge. With light falling on all photocells, the potentiometer is so adjusted that no voltage appears across the relay coil. When *S* blocks a photocell, current flows through the coil of the polarized magnetic latching relay.[1] If either photocell P_1 or P_3 (which are located opposite to each other) is blocked, the current flows to the left through the coil in Fig. 2-a, causing the contact in Fig. 2-b to move (*e.g.* to the right and to remain latched in that position after the block is removed). The relay contact will remain on that side until *S* moves to an adjacent photocell, *i.e.* either to P_2 or P_4. Then current will flow to the right through the coil, moving the contact to the opposite side, where it will remain until *S* again interrupts P_1 or P_3. Each time the contact shifts position, the capacitor in Fig. 2-b charges or discharges through the counter coil and registers through counter E. This occurs only when *S* moves from one photocell to an adjacent one, a distance of about 9 or 10 in. Repeated interruptions of the same photocell, produced by rearing and other non-locomotor behavior, do not yield additional counts.

---

[1] These relays (Sigma #JOZT-3200) are sold by the Universal Relay Corporation, 42 White St., New York, N.Y., or (Sigma #72 AOZ-400TG-TCP) by the Newark Electronics, 500 N. Pulaski Ave., Chicago, Illinois.

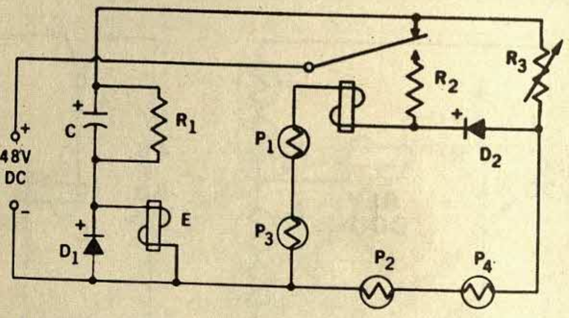


FIG. 3. RECORDING CIRCUIT USING SENSITIVE CONVENTIONAL RELAY
D_1 is for arc suppression. C = 100 to 200 mfd. (typically). R_2 = 5,000 to 20,000 Ω (not critical). R_3 is so selected or adjusted that the relay will not close when photocells are lighted. D_2 and D_1 are silicon top hat or epoxy rectifiers, rated at 100-v. or higher. R_1 should be several times the resistance of the counter, and 2-w. or higher. The relay closes and latches when light is removed from P_2 or P_4 and releases when light is removed from P_1 or P_2.

To test this activity recording system, the photocells were interrupted 300,000 times without any failures in counting.

Fig. 3 shows an alternative recording circuit that used a moderately sensitive conventional relay in place of the polarized latching relay.[2] With the relay open, and light falling on all photocells, the current passes through R_3 and then divides, part going through the relay coil but most going through photocells P_2 and P_4. R_3 is so adjusted that the relay does not pull in. When either P_2 or P_4 is darkened, its resistance becomes very high, allowing more of the current through R_3 to pass through the relay coil, pulling in the relay, which latches itself closed through R_2, P_1, and

---

[2] Sigma 5F16000 or Sigma 5F16000-S/SIL from the Universal Relay Corporation.

P_3. When P_1 or P_3 is then darkened, the relay drops out, applying voltage to the electromagnetic counter E, through capacitor and C resistor R_1, which registers a count. The purpose of C and R_1 is to limit the continuous current through the counter coil while allowing the full voltage of the power supply to be applied momentarily each time the relay releases. This circuit registers a count only every other time *S* moves from one photocell to the next.

To increase the life of the 6-w. lamp, we use a series-resistor to decrease it to 95-v.

A SOLID-STATE SHOCK 'SCRAMBLER'

By STEPHEN W. FLAX and WILLIAM W. HAHN, Denver, Colorado

The frequent use of shock grids for a variety of behavioral treatments and test-measures has led to the development of various electronic and electro-mechanical 'scramblers.' In general, these are devices which apply voltage to different bars to prevent a subject (*S*) avoiding shock by contacting a pair of + or − grid bars, or shorting the grid when a bolus lodges between two bars. The usual disadvantages of previous 'scramblers' have been auditory and electrical noise, excessive bulk and complexity of design, mechnical wear, and failure to provide constant current and voltage. The present apparatus utilizes a simple solid-state design, provides a constant current, produces virtually no electrical or mechanical noise, the cost (approximately $17.00) is considerably less than commercially available devices, and a short-circuit between any pair of adjacent bars does not prohibit a shock from being administered.

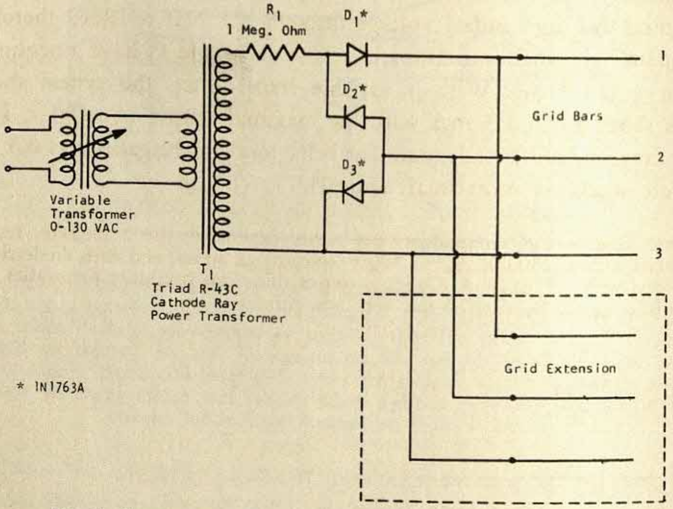


FIG. 1. SCHEMATIC DESIGN FOR SOLID-STATE 'SCRAMBLER'

---

* From The Children's Asthma Research Institute and Hospital, Denver, Colorado. The development of this apparatus was supported in part by USPHS Research Grant HD 02620.

The simplicity of design is accomplished by taking advantage of the phase relationships of line voltage to act as a switching mechanism which channels voltage to different pairs of bars during positive and negative phases of the cycle. Fig. 1 illustrates this technique; during the first half-cycle, shock is administered through grid bars 1 and 2 or 1 and 3; during the second half-cycle, shock is received from the 2-3 combination. The rapid cycling (60/sec.) makes this appear to be a continuous shock. The arrangement of bars can be repeated indefinitely to satisfy requirements for a grid of any dimensions. Theoretically, an *S* could avoid shock by contacting only every fourth bar. In actual practice, however, we have never seen an animal achieve this sophisticated level of learning.

In a typical experiment the output voltage at t_1 is 300–500 v. yielding a grid shock of 0.27–0.45 mA. Relatively constant current is maintained by dropping this high initial voltage through a 1 MΩ resistor;[1] therefore the comparatively small variations in resistance of the *S*s have a negligible effect on grid current. With a variable transformer, the system shown provides shock up to 1.5 mA with the maximal output of 1600 v. from t_1. This range should be adequate for behavioral investigations as 0.5 mA is a severe shock for most small animals.

---

[1] Because high voltage components are rather expensive, the 1 megohm resistor was replaced with 4, 250,000 ¼ w. 500 v. resistors in series, and each diode shown was replaced with 3, 1n1763A diodes in series. It is not necessary to parallel each diode with a resistor because of the very low current.

NOTES AND DISCUSSIONS

APPARENT MOTION OF A VISTA: AN ILLUSION OF PERSPECTIVE

Illusions can usually be characterized as reasonable though erroneous responses by the perceptual system to a pattern of stimulation. An outstanding example is Ames's rotating trapezoidal window, which is perceived as a waggling rectangle.[1] The visual information, though consistent with both the rotation and the waggling, could result in the perception of rotation only if the window is perceived as the trapezoid constructed by Ames. The waggling, however, is consistent for the perspective information that would be obtained from a rectangular window.[2] The view implied is that the percept permits the observer to live in a "real" world that is a best bet about sensory information in that it is consistent with both present and past information. Such a view, which is the common one of many psychologists, colored and directed my reaction to an unusual illusion that I discovered at the Hellbrunn Castle near Salzburg, where I stopped to enjoy some of the pleasures of tourism on the way to the Moscow International Congress of Psychology of 1966.

A *trompe l'oeil* mural, done in about 1618 by Arsenio Mascagni, is the outstanding decoration of the Hellbrunn Banquet Room or *Festsaal*.[3] The mural covers an entire wall, from floor to ceiling, and is about 10 meters wide and 5 meters high. It shows two colonades in linear perspective, with apparently life-sized human figures visible at various points. There is marked aerial perspective for the more distant columns and arches, and the texture of the apparent ground level of the vista is a continuation of the *Festsaal* floor. Rather than claiming to have discovered the illusion, which will be described in a moment, I should admit that our guide pointed it out to me and to the other tourists. His manner indicated that this was one of the well-rehearsed parts of the description of the wonders of Hellbrunn, which must have been presented by guides since the castle was opened to the public in 1922. Furthermore, the illusion is so pronounced that it must have been noticed when the mural was painted, early in the 17th century.

---

[1] Adelbert Ames, Visual perception and the rotating trapezoidal window. *Psychol. Monog.*, 1951, (No. 324). The paper reported here was sponsored by the Aerospace Medical Research Laboratories, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, under Contract No. F 33615-67-C-1280 with Antioch College, Yellow Springs, Ohio. This paper is identified by Aerospace Medical Research Laboratories as AMRL-TR-67-143.

[2] C. H. Graham, On some aspects of real and apparent visual movement, *J. opt. Soc. Amer.*, 53, 1963, 1019–1025.

[3] Mascagni was a Florentine monk in the service of Archbishop Marcus Sitticus Graf Hohenems, who built Hellbrunn Castle. The mural is typically baroque, and similar murals are probably present in many places. This one is illustrated as the second color plate, photographed by Carl Pospesch and titled *Festsaal*, in J. Pupp, *Hellbrunn*, Salzburg, Austria, Salzburger Verlag für Wirtschaft und Kultur, 1965.

The illusion will be described with the help of the perspective drawing in Fig. 1, and can also be seen in that drawing though it is less dramatic and somewhat ambiguous when compared with its appearance in a large mural. Imagine Fig. 1 covering an entire wall of a large room, and imagine yourself walking parallel to this wall. The front or near faces will then appear to be large blocks fixed in the surface of the wall. The illusion occurs as you move: The apparent far ends of the blocks, the ground on which they stand (which appears as a continuation of the floor of the room), and the entire vista will seem to turn in a direction opposite to your motion. Space thus twists to accommodate the continuously changing perspective required by an observer's succesive positions.

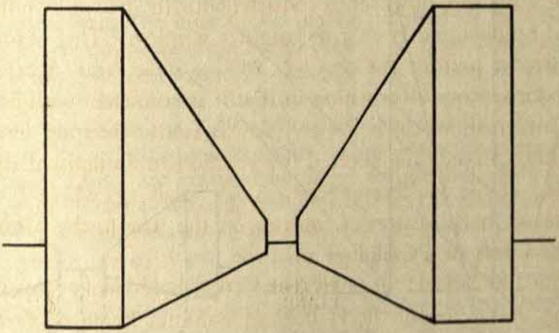


FIG. 1. DRAWING OF A VISTA IN LINEAR PERSPECTIVE
To see the illusion best, the drawing or a transparency of it should be enlarged by projection onto a wall, with the base of the front faces at or near floor level. As one walks by the wall the apparent far ends of the blocks will appear to move in the opposite direction, twisting the vista. The effect can also be seen with any photograph or slide with strong linear perspective.

This illusion can readily be obtained with Fig. 1, if one projects it on a wall or large screen. An alternate illusion frequently occurs with the smaller printed version of Fig. 1, in which the far ends of the blocks and the principal vanishing point are perceived as fixed while the front ends of the blocks turn in the *same* direction as the *O*. To see either illusion most effectively in Fig. 1, one should move one's head a meter or more to the left and to the right while keeping the drawing's lines in clear focus. Monocular regard is not essential.

It is notable that the perceived motion does not involve motion parallax. In fact, the illusion may be thought of as a perceptual world created to accommodate a changing pattern of linear perspective in which the expected motion parallax does not occur

The illusion is related to the familiar effect that the eyes of a well-painted portrait seem to follow the observer.[4] Perspective paintings of familiar objects, such as the recruiting posters of Uncle Sam pointing his finger and saying, "I want you," show a similar effect, and the finger is seen pointing at one for almost all viewing positions. These instances of perspective "following" usually produce three-dimensional perceptual worlds that are not only possible but sensible, and consistent with other experience.

The special effect of the illusion illustrated in Fig. 1, especially when the drawing is large, is that as one walks by it one perceives a physically impossible world. In Mascagni's mural at the Hellbrunn, one sees a moving and twisting vista as one walks by, and the entire space is warped to maintain perspective. In that case the illusion may be enhanced by the additional cues for depth from the aerial perspective and texture grain.

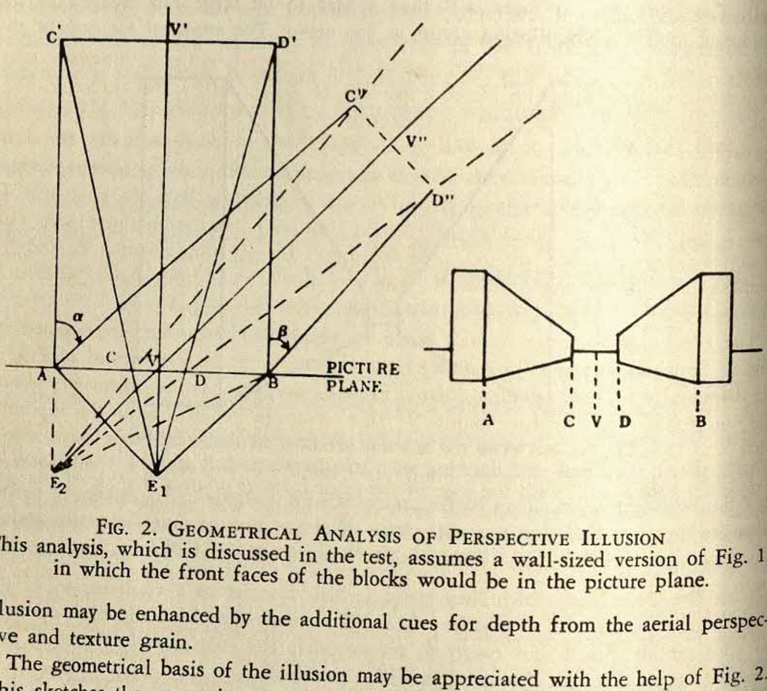


FIG. 2. GEOMETRICAL ANALYSIS OF PERSPECTIVE ILLUSION
This analysis, which is discussed in the test, assumes a wall-sized version of Fig. 1, in which the front faces of the blocks would be in the picture plane.

The geometrical basis of the illusion may be appreciated with the help of Fig. 2. This sketches the approximate situation of the mural in the *Festsaal* at Hellbrunn, or of a wall-sized version of Fig. 1. Points A C V D B are in the picture-plane, as viewed from above, and the eye is in the "correct" position for viewing at E_1. Points C and D are then perceived as projections from the picture-plane to "real" objects C′ and D′. In other words, with the eye at E_1 the mural is seen as representing a depth from the picture-plane backward toward another plane C′D′. The vanishing point, V, is on the line E_1VV′, in which V′ is in the same plane as C′D′. In terms of Fig. 2, the inside edge of the left wall seen from above is perceived to be projected under line AC′, and the right wall is correspondingly under BD′. The illusion occurs when the observer moves away from one viewing point to any other point, and is illustrated for a movement to Point E_2. A line E_2VV'' is the new line through the vanishing point, and gives the new orientation of perspective as perceived by the viewer. Setting $VV'' = VV'$, and assuming that the rear plane will be perpendicular to the line from the observer to the vanishing point (E_2VV'' in the rotated position), the entire perspective is seen to shift with the move to E_2. In the new perspective, the walls have unequal projections as apparently distal stimuli ($AC'' \neq BD''$). Since Points A and B in the picture-plane are also viewed as being in the same plane in the projected real world, the effect is to change the orientation of the colonades or walls with respect to the front edge represented by A and B. The projected rotation of the left wall through Angle α is greater than the projected rotation of the right wall through Angle β, and the distortion of visual space in this illusion is, therefore, nonlinear.

---

[4] E. H. Gombrich, *Art and Illusion: A Study in the Psychology of Pictoral Representation*, 1960, 276.

Fig. 2 was drawn on the assumption of a large mural, with the observer relatively close to the wall. The illusion that can be perceived by the reader in Fig. 1 is somewhat different, because the picture-plane is in front of the near edge of the depicted blocks. Like most perspective drawings, Fig. 1 is seen as a view-through-a-window of a more distant vista. An analysis of the change in perspective when the observer moves is shown in Fig. 3. This can be interpreted in the same way as Fig. 2, with all of the stimulus-information coming from the picture-plane. The projected points in space are indicated by primes for viewing position E_1 and by double-primes for viewing position E_2. The view in Fig. 3 is from above, with each line representing a plane intersecting the surface of the printed page.

The effect shown in Fig. 3 will obviously be weaker than that diagrammed in Fig. 2. Both the near and far planes of the objects in the vista projected in Fig. 3 are displaced when one moves, *e.g.* from E_1 to E_2, and there is less asymmetry about the line through the vanishing point, V. This results in less change in the apparent vista, and there is also less expected motion parallax. Much of the same basic effect is present, however, including the nonlinear distortion of the vista, and will be seen best by the reader if he moves to maximize the distance, E_1E_2. An ambiguity in the illusion offered by Fig. 1 that is not present in the wall-sized mural is easily understood with the help of Fig. 3. The viewer has an option of perceiving either the near or the far plane of the vista as fixed. With the far plane fixed, the illusion is of movement of the front edges of the blocks in the *same* direction as the viewer's movement. My own experience with Fig. 1 is that there is usually some apparent movement of both the near and far planes. The former, in the same direction as head movement, predominates, and the latter, in the opposite direction, is less pronounced.

Two points deserve special emphasis. First, this illusion of perceived movement of a physically stationary field involves active rather than passive movement of the observer to alter the pattern of proximal stimulation. Under conditions of passive movement, *e.g.* of the eyeball or of the seated observer in car or train, apparent movement of objects or of all space is common. I know, however, of no cases, other than the present illusion, where comparably strong disruptions of space constancy follow active movement. Secondly, the illusion involves movement and distortion of perceived space, presumably because of the importance of linear perspective as a cue for depth perception.[5] Rather than give up the perception

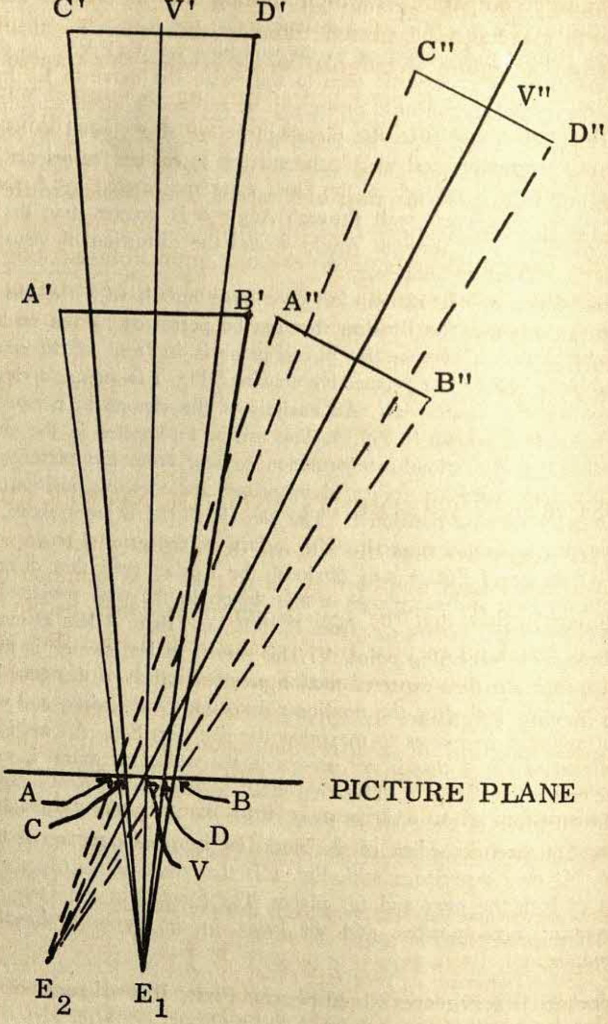


FIG. 3. GEOMETRICAL ANALYSIS OF PERSPECTIVE ILLUSION
This analysis, which is discussed in the text, shows how the illusion occurs for smaller perspective drawings, including Fig. 1 as printed.

---

[5] The cues from aerial perspective and texture in the *Festsaal* mural are probably less significant, because the illusion is strong in wall-sized projections of Fig. 1, in

of depth that would account for the linear perspective, one "creates" an impossible perceptual world that could occur only in a novel physical world. The most nearly comparable familiar illusion may be Ames' trapezoidal window with a metal pipe placed through one pane. To maintain the perception of a waggling rectangular window when this trapezoidal window is rotated, the pipe is perceived as bending in unusual ways. However, in this perception of the trapezoidal window, the dimensions of space remain properly fixed and constant, and the bar takes on a living, snake-like, and in that sense "possible" aspect. The illusion in the *Festsaal* demonstrates an even stronger role for linear perspective: it is a "better bet" for the perceptual system to form an impossibly twisting vista in perspective rather than to ignore perspective and see the mural or painting as it is, an arrangement of colors, textures, edges, and lines on a flat unmoving surface.

Antioch College

HARRY J. JERISON

DISTORTION GRADIENTS AND THE NECKER CUBE

Fisher has suggested that the illusory distortions which appear in the Müller-Lyer and Ponzo figures may be related by a common principle.[1] He suggested further that the spatial distortions which become apparent in figures of this kind may be truly 'geometrical' in that they are evident in any appropriate two-dimensional figure which is not necessarily seen either as having extension in apparent depth, or interpreted as a two-dimensional projection of a three-dimensional structure. The principle invoked in relating the Müller-Lyer and Ponzo illusions implies that predictable distortions should appear in lines and figures whch are located within the framework of inclined lines. The general nature of such distor-

---

which these cues are omitted. It is very likely that the absolute size of the perspective drawing has an important role which I am presently unable to define precisely. The possible significance of size, *per se,* is suggested by Figs. 2 and 3. A larger drawing would be viewed as having its near face closer to or at the picture-plane. Hence the shifting perspective as *O* moves would be accompanied by fixed near points (*e.g.* V) and moving far points (*e.g.* V′). The distortion of space would, therefore, be emphasized.

* Received for publication March 17, 1967. The research reported in this paper was supported by Research Grant No. B/H/139 from the Science Research Council, and No. 70/GEN/9617 from the Ministry of Defence.

[1] G. H. Fisher, A common principle relating the Müller-Lyer and Ponzo illusions, this JOURNAL, 81, 1968, 000-000; F. C. Müller-Lyer, Optische Urteilstäuschungen, *Arch. f. Physiol.,* Suppl. Bd., 1889, 263-270; Mario Ponzo, Rapports entre quelque illusions visuelles de contraste angulaire et l'appréciation de grandeur des astres à l'horizon, *Arch.Ital.d. Biol.,* 58, 1912, 327-329.

tions is illustrated in Fig. 1. This figure consists simply of the Ponzo illusion, added to which are further horizontal lines. It shows very clearly that a gradient of apparent-distortion appears in figures enclosed within inclined lines. Each successively lower line appears to be slightly shorter than its upper neighbour until the lines appear in positions below the locus and influence of the oblique lines.

Many theories have been proposed in the attempt to relate illusory spatial distortions to apparent depth. Those due to Thiéry and Woodworth invoke the principles of perspective to account for them,[2] while the theory proposed by Tausch and developed by Gregory implies that they may be explained according to the inappropriate operation of the mechanism underlying size-constancy.[3] All theories of this kind imply that illusory figures are either seen as having extension in apparent depth, or are typical two-dimensional projections of three-dimensional scenes. Fisher has suggested that these views may be mistaken and that depth features, as such, are irrelevant to explanation of the distortions which appear in the geometrical illusions.[4] A more parsimonious theory than those which invoke the influence of further perceptual mechanisms, knowledge of the modes of operation of which is incompletely understood, would be one capable of understanding all illusions under a common principle to which further assumptions were unnecessary. The 'distortion-gradient' theory is one which appears capable of explaining many examples of the geometrical illusions in a consistent and systematic way. A critical test of the theory would be that of considering the extent to which it is applicable to figures which are undoubtedly seen as having extension in apparent depth. If the theory proved capable of explaining the distortions seen in such figures without appeal to the influence of the depth figures contained within them, it would clearly be advanced a stage further.

A convenient figure to consider in the context of this theory is the Necker cube,[5] illustrated in Fig. 2(a). This is a skeleton figure which, upon initial observation, usually tends to assume the form illustrated in Fig. 2(b). Subsequent to prolonged observation, however, this figure ap-

---

[2] Armand Thiéry, Über geometrisch-optische Täuschungen, *Phil.Stud.*, 12, 1896, 67-125; R. S. Woodworth, *Experimental Psychology*, 1938, ———.

[3] R. Tausch, Optische Täuschungen als artifixielle Effekte der Festaltungprozesse von Grössen- und Formenkonstanz in der natürlichen Raumwahrnehmung, *Psychol. Forsch.*, 24, 1954, 299-349; R. L. Gregory, Distortion of visual space as inappropriate constancy scaling, *Nature*, 199, 1963, 678-680.

[4] Fisher, Illusions and size-constancy, this JOURNAL, (in press)

[5] L. A. Necker, Observations on some remarkable phenomena seen in Switzerland; and an optical phenomenon which occurs on viewing of a crystal or geometrical solid, *Phil.Mag.*, *3 ser.*, 1, 1832, 329-343.

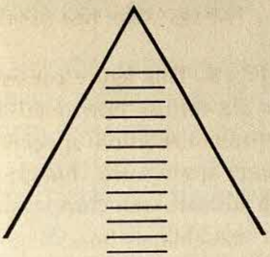
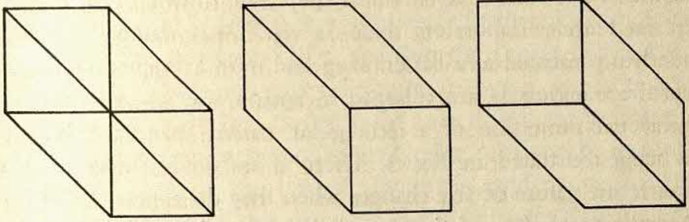


FIG. 1. THE GRADIENT OF APPARENT DISTORTION SEEN IN A NEW VERSION OF THE PONZO ILLUSION



a. b. c.

FIG. 2. THE NECKER CUBE AND INDICATIONS OF ITS APPEARANCE IN TWO ASPECTS

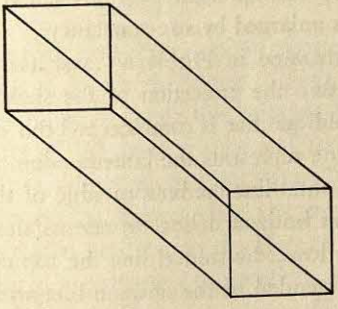
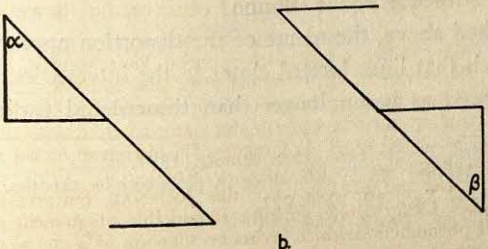


FIG. 3. A RECTANGULAR SKELETON FIGURE



a. b.

FIG. 4. COMPONENT PARTS OF THE NECKER CUBE

pears to reverse in aspect assuming the form indicated in Fig. 2(c). Since the square faces of the figure, orthogonal with the angle of regard, are in different apparent-positions depending upon the aspect in which it is seen, they should undergo appropriate changes in apparent-size. There is some disagreement as to whether such changes do, in fact, appear. Sanford[6] and Hotaopf[7] reported that they do so; Gregory[8] reported being unable to observe them in two-dimensional drawings of the Necker cube, but that they appeared clearly when the figure was drawn in luminous-paint and viewed in ultra-violet light in such a way as to remove the paradoxical influence of the surface upon which they were drawn..

In the author's laboratory some *S*s report size-changes in the Necker cube when presented as a flat drawing and most *S*s appreciate them readily when the transversals are enlarged in such a way as to represent a two-dimensional projection of a rectangular, rather than sube-shaped figure, this being illustrated in Fig. 3. There is no doubt, however, about the characteristic nature of size changes when they do appear. *In each case the apparently nearer face of the figure is judged to be smaller than that which is apparently further away.* The appearance of such differences in figures seen in reduced-cue conditions led Gregory to propose the notion of 'secondary constancy scaling' according to which the apparently further face of the figure is enlarged by size-constancy.

The patterns illustrated in Fig. 4(a) and (b) will be recognized as being contained within the projection of the skeleton cube shown in Fig. 2(a). The longer oblique line is common to both figures. In Fig. 4(a) the upper, horizontal line represents the bottom edge of the upper square face and the lower horizontal line the bottom edge of the lower square face. In Fig. 4(b) the upper horizontal line represents the top edge of the upper square face and the lower horizontal line the top edge of the lower square face. Angles are subtended by the oblique line with the left-hand edge of the upper square face in Fig. 2(a) and with the right-hand edge of the lower square face in Fig. 2(b). Eachof these figures is thus a partially completed version of the Ponzo illusion.

As described above, the nature of the distortion appearing in the Ponzo figure us such that lines located closer to the intersection of lines forming angles are seen as being longer than thoseplaced further away. If this principle is relevant to the distortion which is evident in the Necker cube, the upper horizontal line should appear longer in Fig. 4(a) and the lower horizontal line should appear longer in Fig. 4(b). What happens when the Necker cube reverses in aspect is that certain of its component parts become figural, others being relegated to the background. When the cube tends towards the form of Fig. 2(b), angle α becomes figural, angle β becoming figural when it assumes the form of Fig. 2(c). Hence, when either of the angles, α or β, become figural, the pattern should assume a characteristic aspect, whereupon predictable changes in size should be evident. When α is figural the lower face is apparently nearer than the upper face and when β is figural it is apparently further. In each of these aspects, however, the horizontal line, and thus the square face associated with ita, is the one which undergoes distortion in such a way as to appear longer. But this is also the line closest to the apex of the angle. Hence, the size changes which become apparent in the Necker cube may be understood as being special cases of the spatial distortions evident in the Ponzo figure. Thus, application of this principle to the apparent changes in size of the faces of the Necker cube upon reversal yields a consistent explanation of its appearance in each aspect also, without recourse to depth factors.

---

[6] E. C. Sanford, *Psychologie experimentale*, 1897, ——.

[7] W. H. N. Hotopf, Some difficulties in perspectivist theories of visual illusions. *Bull Brit. Psychol. Soc.*, 19, 1966, 922.

[8] Gregory, *Eye and Brain*, 1966, 156.

SUMMARY

The gradient of apparent distortion seen in figures contained within patterns of inclined lines, of which the Ponzo illusion is a special case, is described. An attempt is made to apply this princple to the dfferences in apparent size of the faces of the Necker cube and to the changes in size evident upon reversal. It is suggested that the Ponzo illusion and the two alternative figural aspects of the Necker cube are similar. Application of the principle of the distortion-gradient to the Necker cube in each of its two aspects appears to yield a consistent explanation of its appearance.

University of Newcastle Upon Tyne GERALD H. FISHER

PREFERENCES FOR BODILY PARTS FOLLOWING MASTECTOMY

In a previous study, we reported that normal adults of varying sex, age, and socio-economic status were highly consistent in their rankings of the value of selected bodily parts.[1]

The purpose of the present investigation was to assess the validity of the previous results by ascertaining preferences for the same parts in a population of subjects which had sustained loss of an appendage.

The *S*s were 538 women with unilateral mastectomies (273 right, 265 left; mean age = 54 yr.; range 30–79 yr.) performed from less than 2 to more than 9 yr. earlier, and 52 women with bilateral mastectomies (mean age = 55 yr.; range 40–69 yr.) performed 2 to 32 yr. earlier.

The questionary consisted of a single sheet with space for recording *S*'s age, education, occupation (if housewife, previous occupation), site of mastectomy (right, left, or both breasts) and date(s) of mastectomy. A list of bodily parts with a space next to each for the rank completed the form. The instructions given were like those used in the earlier study. The questionary was self-administered and returned by mail. Replies were received from most areas of the country.

Fig. 1 presents the mean ranks of parts as a function of age for all unilateral cases of mastectomy. As can readily be seen, there is a high degree of consistency among all age-groups. In fact, the concordance in the present study appears greater than that obtained previously from a comparable group of women without amputation. The mean ranks for the breast alone, plotted as a function of age, for women without mastectomies (data from the previous study), for *S*s with unilateral mastectomies, and for *S*s with bilateral mastectomies, are shown in Fig. 2. It can be seen that, compared to normals, mastectomy reduces somewhat the relative value of the breast, particularly for *S*s with bilateral amputation. It should be noted, however, that the breast, almost without exception, maintains its rank among all other parts for groups of different age, with or without mastectomies. It also is clear that unilateral mastectomy reduces variability across the various ages and also tends to reverse the trend found in nor-

---

* This work was supported by Grant RD-1495-M from the Vocational Rehabilitation Administration. The authors are indebted to Mrs. Terese Lasser for permitting distribution of a questionary to members of Reach to Recovery.

[1] Sidney Weinstein, E. A. Sersen, Larry Fisher, and R. J. Vetter, Preferences for bodily parts as a function of sex, age, and socio-economic status, this JOURNAL, 77, 1964, 291–294.

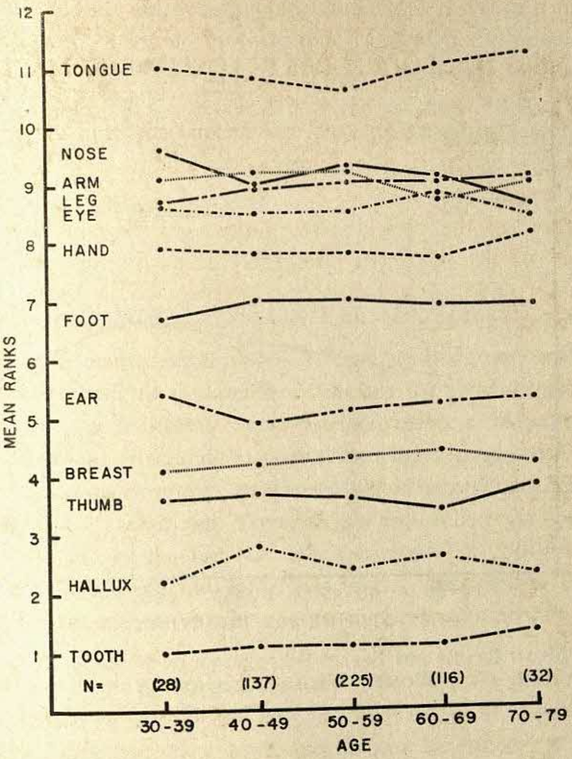


FIG. 1. BODILY PART-PREFERENCES AFTER UNILATERAL MASTECTOMY AS A FUNCTION OF AGE

mals, namely, that the breast tends to become slightly less important with age.

The data were divided according to right or left amputation, socio-economic level (as defined in the previous study), age at time of reporting, and years since operation, as specified in Table I. Degree of consistency between these groups was determined by rank-ordering the mean values for parts. These ranks then were compared using Spearman's rank-order correlation (rho) where two groups were compared, and by Kendell's coefficient of concordance (W) where three or more groups were compared. The results indicated that a high degree of concordance ($p < 0.001$) among various groups regardless of the basis of classification. Specifically, all coefficients ranged from 0.977–0.996. The correlation between the rank-order preferences in all normal women and all unilateral mastectomy cases also was highly significant (rho $= 0.969$; $p < 0.001$).

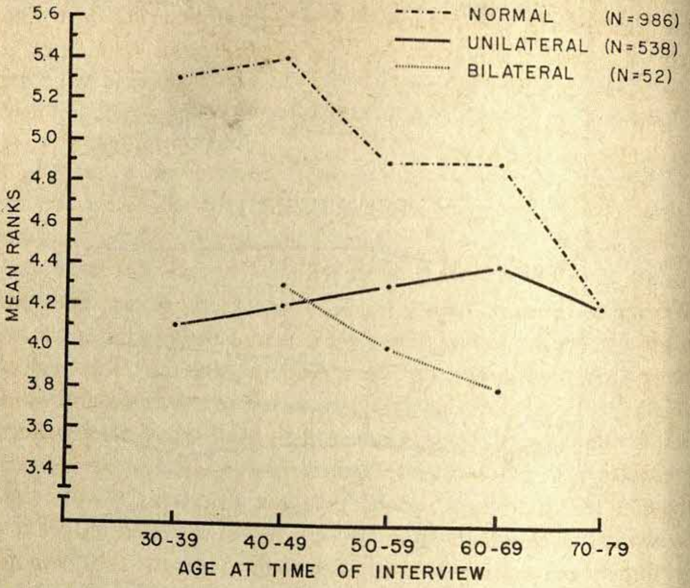


FIG. 2. MEAN RANKS FOR BREAST-PREFERENCE IN NORMAL WOMEN AND WOMEN WITH MASTECTOMIES

TABLE I

DISTRIBUTION OF UNILATERAL MASTECTOMY CASES ACCORDING TO AGE AND TIME SINCE OPERATION

Time Since Operation (Years)	Age (yr.)					Total
	30–39	40–49	50–59	60–69	70–79	
<2	4	33	32	13	4	86
2	12	18	20	9	5	64
3	6	13	20	9	3	51
4	1	8	20	4	0	33
5	0	8	13	2	2	25
6	0	10	10	6	0	26
7	0	12	20	6	2	40
8	2	17	37	17	7	80
9	2	6	16	10	2	36
>9	1	12	37	40	7	97
Total	28	137	225	116	32	538

It is evident, then, that mastectomy does little to modify the relative value of the bodily parts we have studied. Although the subjective value of the breast is slightly reduced following amputation, it continues to maintain the same relative position and similar mean rank.

New York Medical College SIDNEY WEINSTEIN
ROBERT J. VETTER
Jewish Hospital of Brooklyn EUGENE A. SERSEN

THE FORM OF LOGICAL SOLUTIONS

The aim of heuristic principles is to find methods and rules to assist discovery and invention.[1] Golann has reviewed the psychological study of creativity from four points of view: products, process, measurement, and personality.[2] The present analysis is restricted to *products*, and consists of the result of a rather simple linguistic analysis of one formal characteristic of 'creative' problem-solutions. Vygotsky has argued for the independence of thought and language, but he does not deny that people can think about words.[3] A heuristic question is whether there might be a set of words that are particularly worthy of thought, when the goal is to discover or invent.

A well-known formal characteristic of many creative problems is that they often consist of taking two things in a new relation to each other. How many relations exist in which things may be taken? A potentially useful list of such relations follows from a few successive simplifying assumptions. Assume that each such relation can be stated in an English sentence. It follows that the set of all randomly generated English sentences must include a subset of relational sentences. The vocabulary of normal English is, however, very large. Reduce the set of words to be considered to the 850 words of Ogden's *Basic English*.[4] Basic English is a short form of English in which the needs of everyday life can be conveyed; the recent publication of *The Basic Dictionary of Science* also shows that the normal words that describe scientific objects and concepts can be translated into Basic English without a loss of meaning and with a gain in clarity.[5] Assume that the set of all randomly generated Basic English sentences would include a subset of relational sentences. Such sentences can take a variety of forms but common to them all is that one thing is taken in some relation to another thing; the elementary form of such sentences is "Take A (one thing) in some relation to B (another thing)." The full set of allowable relations in such an elementary sentence is shown in Table I.

---

[1] G. Polya, *How to Solve It*, 1957, 112.

[2] S. E. Golann, Psychological study of creativity, *Psychol. Bull.*, 60, 1963, 548–565.

[3] L. S. Vygotsky, *Thought and Language*, 1962, 51.

[4] C. K. Ogden, *The System of Basic English*, 1934, 4.

[5] E. C. Graham, *The Basic Dictionary of Science*, 1965, 1-54, *passim*.

The splendid variety of human inventions and discoveries is found in the products of men engaged in problem solving of the highest order; what cognitive steps that each made can not be known. But surely they were not equipped with a set of elementary sentences that could be used as templates to guide them to the form of solutions. Whatever else the Periodic Table is, it is also the elementary sentences "Take A (one element) after B (another element)," on the basis of atomic number, and "Take A under B," on the basis of similar chemical properties. Similarly, the discovery of Neptune by Adams and Leverrier was also "Take A (the mass and orbit of an unknown body deforming the orbit of Uranus) from B (the orbit of Uranus)." And similarly, installing a fountain in one's garden to mask the noise of a nearby highway is also "Take A (the splashing of water) over B (the unwanted noise)."

TABLE I

THE RELATION-WORDS

about
across
after
against
among
and
as
at
because
before
between
but
by
down
for
from
if
in
near
out
now
of
off
on
opposite
or
out
over
round
still
so
then
though
through
till
to
under
up
when
where
while
with

The templates exist to be filled, and it may be possible to work backwards from them to problem-solutions. Even with them, much is required of a problem-solver: the choice of the goal, of the things to be related, and of the templates to be used. Nonetheless, the templates may have heuristic value in directing attention to the form of solutions.

Durham, North Carolina
HERBERT F. CROVITZ

GANZ'S HYPOTHESIS ON FIGURAL AFTER-EFFECTS

Recently, Ganz has proposed a new hypothesis to explain figural aftereffects.[1] The major assumption of this hypothesis is that the *after-image* of the inspection-figure (*I*) interacts with the test-figure (*T*) as in a simultaneous geometric illusion. In the examples given by Ganz,[2] the *T*-figure is displaced away from the *I*-figure when the *I*- and *T*-figures are presented simultaneously, as well as when the *T*-figure is presented after the *I*-figure has been removed. That is, the contours of the *I*-figure always repel the contours of the *T*-figure whether the two are presented simultaneously or successively.

---

[1] Leo Ganz, Mechanism of the figural aftereffects, *Psychol. Rev.*, 73, 1966, 128–

Ganz supplies inferential support for his hypothesis, including evidence that after-images and after-effects show very similar decay rates. A more direct test is, however, available. Are there any *I*- and *T*-figures which, when shown simultaneously, produce an illusion which is directly *opposite* to the illusion created when the *T*-figure follows, after some fixation, the *I*-figure in time? If there are, then it is difficult to see how an after-image can provide an adequate explanation of these after-effects.

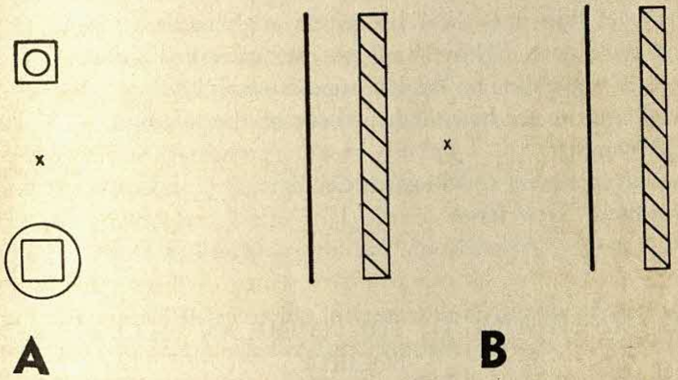


FIG. 1. STIMULI WHICH, WHEN PRESENTED SIMULTANEOUSLY, INDUCE AN ILLUSION OPPOSITE TO THAT INDUCED WHEN THE STIMULI ARE PRESENTED SUCCESSIVELY

In A, the inspection-figure consists of the circles, the test-figure consists of the squares; in B, the inspection-figure consists of the hatched bars, the test-figure consists of the solid lines.

In fact, there are such examples, two of which are given in Fig. 1:A, showing a change in size and B showing a displacement in space. For the exact dimensions of these stimuli, the reader is referred to the original publications. In both sets of figures, the "X" is used as a fixation-point.

Example 1 (Fig. 1-A). These figures were used by George,[3] and sub-

---

150; Is the figural aftereffect an *after*effect? A review of its intensity, onset, decay, and transfer characteristics, *Psychol. Bull.*, 66, 1966, 151–165.

[2] Ganz, *op. cit.*, *Psychol. Rev.*, 73, 1966, 128–150, Figs. 2–4.

[3] F. H. George, On the figural aftereffect, *Quart. J. exp. Psychol.*, 5, 1953, 128–135.

sequently by Spitz and Blackman.[4] When the circles are used as *I*-figures, which are then removed and replaced by the squares (*T*-figures), the bottom square looks *smaller* than the top square, although both squares are objectively equal. When all figures are presented simultaneously, as shown here, the bottom square looks *larger* than the top square.

Example 2 (Fig. 1-B). These figures were adapted from Wertheimer[5] and changed to resemble those shown in Fig. 1-B by Spitz and Lipman.[6] When the cross-hatched bars are used as *I*-figures, then replaced by the solid line *T*-figures, the *T*-line on the left appears to be *farther* from the fixation-point then does the *T*-line on the right, although they are both actually equidistant. When all figures are presented simultaneously, as shown here, the *T*-line on the left appears *close*r to the fixation-point.

It is difficult to see how the hypothesis of the after-image can explain these phenomena.

Johnstone Training and Research Center
Bordentown, New Jersey
HERMAN H. SPITZ

THE 1966 MEETING OF THE AMERICAN PHILOSOPHICAL SOCIETY

The Annual General Meeting of the American Philosophical Society was held in the historic Hall of the Society (built by Benjamin Franklin) in Philadelphia on April 20, 21, and 22, 1967.

Under nomination of members in Class III (Social Sciences) the Society this year elected one psychologist to membership. The new member is Dr. Lee J. Cronbach, Professor of Education and Psychology at Stanford University.

The program of addresses at the meetings included a larger number of papers in psychology than is often the case. These papers were: "Laws of learning relevant to its biological basis," by Neal E. Miller; "Is there a biochemical correlate to the engram?" by Holger Hydén; "Analysis of memory storage enhancement and transfer," by James L. McGaugh; "Memory and learning from the standpoint of computer model building," by Richard B.

---

[4] H. H. Spitz and L. S. Blackman, A comparison of mental retardates and normals on visual figural aftereffects and reversible figures, *J. abnorm. soc. Psychol.*, 58, 1959, 105–110.

[5] Michael Wertheimer, Figural aftereffects as a measure of metabolic efficiency, *J. Pers.*, 24, 1955, 56–73.

[6] H. H. Spitz and R. S. Lipman, Reliability and intercorrelation of individual differences on visual and kinesthetic figural aftereffects, *Percept. mot. Skills*, 10, 1960, 159–166.

Roberts; and "Slow electrical responses of the human cortex," by Hallowell Davis.

National Geographic Society LEONARD CARMICHAEL

THE 1967 MEETING OF THE NATIONAL ACADEMY OF SCIENCES

The National Academy of Sciences held its 104th Annual Meeting at the Academy building in Washington on April 24, 25, and 26, 1967.

Under direct nomination of the section on Psychology, Dr. James J. Gibson, Professor and Chairman of the Department of Psychology at Cornell University, and James Olds, Professor of Psychology at the University of Michigan were elected to membership. In cooperation with other sections, Dr. Herbert A. Simon, Professor and Associate Dean in the graduate school of Industrial Administration in the Carnegie Institute of Technology, and Dr. Walle J. H. Nauta, Professor of Neuroanatomy in the Department of Psychology of the Massachusetts Institute of Technology were elected.

The following papers of special interest to psychologists were presented: "A quantitative study of pain and its reduction through hypnotic suggestion," by Ernest R. Hilgard; "Roles of intellectual abilities in the learning of concepts," by J. P. Guilford, J. L. Dunham, and R. Hoepfner; "A polygenic theory of schizophrenia," by Irving I. Gottesman and James Shields; "Estimation of the limits of heritability of traits by comparison of monozygotic and dizygotic twins," by Arthur R. Jensen: "A 'Try-Simplest Cases' approach to the heredity-poverty-crime problem," by W. Shockley: "Racial differences in intelligence: still an open question?" by Wallace A. Kennedy; "Chemical signaling between polyphemus Moths and between Moths and Host Plant," by Carroll M. Williams and Lynn M. Riddiford.

Leonard Carmichael was the speaker at the annual banquet of the Academy. His topic was "Joseph Henry and the National Academy of Sciences."

National Geographic Society LEONARD CARMICHAEL

SIXTY-THIRD ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The sixty-third annual meeting of the Society of Experimental Psychologists was held at the University of Illinois on March 31 and April 1,

1967. Lloyd G. Humphreys, Chairman of the Society for 1966–67, served as Chairman of the meeting.

Members and fellows present were: Anderson, Bray, Brodgen, Brown, Chapanis, Dallenbach, Duncan, Egan, Eriksen, Estes, E. J. Gibson, J. J. Gibson, Grant, Grice, Guilford, Hake, Hirsh, Humphreys, Kappauf, Kendler, Kimble, Lanier, Luce, McCleary, McGill, Melton, Meyer, Morgan, Olds, Postman, Riggs, Solomon, Stellar, Thomas, Torgerson, Underwood, Wickens.

Richard C. Atkinson, Eckhard H. Hess, James J. Jenkins, John I. Lacey, Frank A. Logan, and Mark R. Rosenzweig were elected to membership in the Society, bringing the total to 68 members and 34 fellows.

To mark the hundredth anniversary of the birth of the founder of the Society, Titchener, the history of the Society was reviewed by Boring in absentia by means of a taped record due to his illness; and Dallenbach announced the establishment of a Medal for outstanding research in Sensory Psychology.

Reports of research in progress were presented in scientific sessions on March 31 and April 1.

The Warren Medal for 1967 was awarded to Eliot Stellar "in recognition of his long-term, broad-scale experimental attack upon physiological factors involved in hunger and thirst and his theoretical formulation of relations between the hypothalamus and motivation."

The Society accepted the invitation to meet at The Rockefeller University, New York City, in 1968. Carl Pfaffmann was elected Chairman for 1967–68.

University of California, Berkeley LEO POSTMAN

THE FIFTY-NINTH ANNUAL MEETING OF THE SOUTHERN SOCIETY FOR PHILOSOPHY AND PSYCHOLOGY

The fifty-ninth Annual Meeting of the Southern Society for Philosophy and Psychology was held on March 23–25, 1967, at the Hotel Roanoke in Roanoke, Virginia. Hosts were the Departments of Philosophy and Psychology at Washington and Lee University, Virginia Military Institute, and the University of Virginia. Approximately 325 members, friends, and guests attended.

The program consisted of one symposium in philosophy and three symposia in psychology, a general discussion in psychology on the report of the commission on composition of council of the A.P.A., and two joint sessions. Altogether there were 33 papers in philosophy and 40 papers in psychology. As usual, a social hour and business meeting were also included in the program.

At the first of the two joint sessions, Edward G. Ballard, Tulane University, delivered the Presidential Address entitled "Perceptual distance." At the second joint session Professor Charles Hartshorne, University of Texas, delivered an invited address entitled "Psychology and the unity of knowledge." Discussants of this address were Carroll Pratt, Princeton University, and William S. Weedon, University of Virginia.

Two Junior Awards were announced: one for the paper in philosophy presented by George B. Thomas, University of Virginia, entitled "He could not have chosen otherwise," and the second for a paper in psychology presented by Henry B. Van Twyver, University of Florida, on "GSR conditioning in immobilized cats." These awards are given annually for the papers of highest merit presented by members who have not yet received the doctoral degree or who have received it within the last five years.

At the annual business meeting it was announced that the Council had accepted the invitation of the University of Miami to hold the sixty-first Annual Meeting in Miami, Florida, April 3–5, 1969.

The death of Professor Max F. Meyer was noted with sorrow and a memorial was read.

Twenty-five philosophers and 38 psychologists were elected to full or associate membership.

The following officers were elected to serve during 1967–68: Earl A. Alluisi, University of Louisville, President; E. M. Adams, University of North Carolina, President-elect; Edward G. Ballard, Tulane University, Past President. Douglas Browning, University of Miami, continues to serve as Secretary. The term of the Treasurer, Dwight Van de Vate, Jr., University of Tennessee, was extended for one year. John W. Davis, University of Tennessee, and Milton H. Hodge, University of Georgia, were elected to the Council for three-year terms as representatives of philosophy and psychology, respectively. Other continuing members of the Council are Malcolm D. Arnoult, William T. Blackstone, Lewis Hammond, Clyde E. Noble, and Herbert C. Sanborn.

The sixtieth Annual Meeting of the Southern Society for Philosophy and Psychology will be held on April 11-13, 1968, in Louisville, Kentucky.

University of Miami DOUGLAS BROWNING

THE THIRTY-EIGHTH ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The Eastern Psychological Association held its thirty-eighth Annual Meeting April 6–8, 1967, in Boston, Massachusetts. The Sheraton-Boston Hotel served as headquarters. Total registration at the meeting was 3,370. The membership of the Association at the time of meeting totalled 4,322.

Andrew S. Dibner, chairman of the Committee on Local Arrangements, was assisted by A. George Gitter, J. Edward Conners, Teresa Carterette, Lewis J. Sherman, Herbert J. Hoffman, and Alice L. Palubinskas. The Program Committee, consisting of Donald L. Grant, Chairman, Roger Bibace, and L. Starling Reid, scheduled 515 papers, 9 symposia, 1 round table discussion, 2 special sessions, 3 film sessions, 5 special meetings, and 2 invited addresses.

The 100 paper-reading sessions were distributed in terms of general topics as follows: 15 on physiological psychology; 16 on sensation and perception; 7 on operant behavior; 15 on human-experimental problems; 8 on animal-experimental problems; 7 on developmental and child behavior; 18 on social behavior and personality; 10 on abnormal and clinical psychology; and 4 additional sessions covering statistics, personnel and educational psychology.

Specialized session titles, new to the program this year, were physiological-memory, limbic system, hypothalamus and motivation, food deprivation, feeding and drinking, nystagmus, signal detection, infant behavior, risk-taking, disadvantaged youth, community mental health, and clinical-social psychology.

A round table discussion was conducted on "Stages of development." Brief titles of symposia were "Personality theory and conceptual structure," "Effects of E.D.P.," "Spatial maps and codes," "Effects of early stimulation in childhood," "Approaches to learning difficulties in boys," "Psychotherapy and elderly patients," "Measurement of learning from connected discourse," and "Brief response-produced cues."

James Deese delivered the Presidential Address, "Meaning and change of meaning." Kenneth E. Clark gave an invited address, "Invasion of privacy in the investigation of human behavior." Also, by special invitation, Karl H. Pribram spoke on "The new neurology and the biology of emotion." The Psi Chi invited address "Measuring by matching" was given by S. S. Stevens.

At the Annual Business Session, Eleanor J. Gibson was announced as President-elect and Joseph V. Brady, Morton Deutsch, and Philip Teitelbaum as new Directors of the Association. William W. Cumming was appointed to the office of Executive Secretary to succeed Marvin A. Iverson.

A budget of $27,185 was approved for the forthcoming fiscal year. The next Annual Meeting will be held April 18–20, 1968, in Washington, D.C. with the Sheraton-Park Hotel as meeting headquarters.

Queens College of the City University of New York
MARVIN A. IVERSON

ERRATUM

A line of type was inadvertently dropped from the Book Review Section of the March number of this JOURNAL (Vol. 80, 1967, p. 162) which obscures the reviewer's meaning. The sentences concluding the second paragraph on p. 162, beginning on line 21, should read as follows, with the line dropped being set in brackets: "The discussion inevitably supports the use of concealed methods if it strongly states the reactive effects inherent in direct and open observation. [The extent to which a subject is affected by awareness of observation] depends upon his sensitivities, his interpretation of the observation's purposes, the compellingness of other aspects of his situation, the length of time he has been under observation, and other factors."

K.M.D.

BOOK REVIEWS

Edited by T. A. Ryan, Cornell University

The Causes and Cures of Neurosis. By H. J. EYSENCK and S. RACHMAN. San Diego, Robert R. Knapp, 1965. Pp. xii, 318. $7.95.

8he book is aptly subtitled "An Introduction to Modern Behaviour Therapy Based on Learning Theory and the Principles of Conditioning." It is a textbook providing the theoretical rationale, case illustrations of techniques, and an evaluation of this important new school of clinical work. Behavior therapy started with Wolpe's book in 1958 and Eysenck's christening term in 1959. As an idea and occasional procedure, it can be traced back to Watson and Mary Cover Jones in 1920 and 1924. An illustration of cultural lag! In America, a Skinnerian stream has joined the Wolpian-Eysenkian stream, producing a formidable torrent of behavioristic publication that promises to be one of the most research-generating happenings in clinical psychology. Half of the references in the Eysenck and Rachman book are between 1960 and early 1964, when the book went to press, and there have been many publications since then. Such a challenge would inevitably command a response. Breger and McGaugh (*Psychol. Bull.*, 1965, 63, 338-358) opened with a salvo from the cognitive-perceptual fort which hopefully will lead to greater specification of alternatives and a series of critical experiments.

What is behavior therapy? Eysenck and Rachman view "symptoms" simply as maladaptive habits, not indicators of unconscious complexes. Pavlovian conditioning accounts for the development of the great majority of neurotic disorders characterized by anxiety, obsessions and other dysthymic symptoms. These kinds of symptoms often remit spontaneously but may be preserved from extinction when the person avoids repeated exposure to the uncomfortable stimulus. Another kind of neurotic disorder is caused by a failure in the conditioning of socially desirable habits, exemplified by psychopathic behaviors, eneuresis, and non-anxious hysterical symptoms.

At this point, the authors ingeniously relate neurotic disorders to Eysenck's personality theory and to biological factors. Dysthymic patients tend to be strong introverts; those with conduct problems, extraverts. Neuroticism and introversion-extraversion are seen as growing out of original constitutional propensities. Eysenck's theory postulates that introverts are more subject to excitation and facilitation of processes of remembering and learning, whereas extraverts tend toward greater distractability and readier inhibition of cortical processes. Citing several studies, the authors show that introverts condition more readily than extraverts.

In 11 of the 17 chapters, Eysenck and Rachman survey all known reports of the use of behavior therapy. Techniques are too varied to report in detail. They include Wolpe's systematic desensitization to anxiety-generating stimuli, relaxation and training in self-assertion. The foregoing are largely used with dysthymic, introverted persons with symptoms of anxiety, phobias, obsessions, and compulsions. With undesirable behaviors, often showing low anxiety, the authors cite the effective use of extinction by non-reinforcement, conditioning of new responses, negative practice (massed fatiguing repetition of the undesirable habit), and aversive conditioning.

Drug facilitation of conditioning is sometimes used. The authors provide few illustrations of sociopathic behavior. They show operant conditioning as a promising technique for improving psychotics' behavior. Cases are often treated by more than one technique. The authors also recommend follow-up "booster treatments." Always the focus is on eliminating or changing specific behaviors. The many arrangements for giving immediate feedback to patients demonstrate much flexibility and innovation.

In evaluating a textbook for clinical work, the following criteria should be used: clarity, relevancy, effectiveness and theoretical soundness. Does the book communicate the nature of behavior therapy? A clinician reading the book can get a fairly clear picture of the techniques, but a less clear picture of the cases. Recent applications of "reinforcement therapy" to mental hospital patients indicate that aides and nurses can understand and apply operant conditioning principles. Since behavior modification may be repetitious, boring, and time-consuming, it is important that the approaches be usable by less highly trained people. Clear as the book is, however, it could benefit by more exact reporting of what is done—especially about (a) the identification of the symptom, (b) the choice of treatment-technique, and (c) the context of treatment. For research purposes, it would be important to disentangle behavior-technique *per se* from interpersonal influence.

How relevant is this presentation to problems in clinical work? Many illustrative cases have specific symptoms, *e.g.* writer's cramp, phobias, eneuresis, impotence. Sometimes an individual with pervasive problems appears to be defined as having one or two narrow problems. In the first and most extensively reported case, a young woman with problems of sexual relations is treated through desensitization, using an "injection hierarchy" and a "sanitary pad hierarchy" (p. 69). When she was too troubled to participate in the desensitization process, the session "was restricted to a sort of non-directive, cathartic discussion" (p. 73). This report raises the question: Just what is the effective agent in behavior therapy? Little attention is given to interpersonal problems or to the possibility that the "identified patient" in a family is actually the one with the problem. By its emphasis on observable symptoms of individuals, behavior therapy deemphasizes other aspects of persons—especially the social system and the self-system. A comparison of frequently indexed terms in the Eysenck and Rachman book with a new American survey of therapy research shows an ommision of concepts about (a) social relations, such as "empathy", "communication", "dominance" and "culture", (b) cognitive processes, such as "awareness", "coding", "depersonalization" and "dreams" and (c) emotions, such as "elation", and "ambivalent hostility". On the other hand, the authors more frequently employ names of drugs and diagnostic nomenclature such as "agoraphobia" and "fetishism". The behavior therapy of Eysenck and Rachman is oriented toward clinical problems of individualistic pathology, not to problems of relationship, milieu, decisionmaking, or positive strivings.

How effective are the procedures? This crucial problem in clinical psychology raises many questions about what kinds of techniques used by what kinds of therapeutic agents with what kinds of cases are effective with what kinds of criteria. Of all researchers, Eysenck is aware of the problem. His famous 1952 article pointed out that studies of therapy reported no improvement over the base rates of 65 to 70% for "spontaneous remission". Surveying 10 studies of the effectiveness of behavior therapy, including 3 with control groups, the authors concluded "something in the region of 80 per cent of the patients were apparently cured or markedly improved" (p. 266). It is almost impossible to judge whether the results are importantly different from the natural remission rate until psychologists carry out many finely designed, multiple control group studies on specified clinical problems. At least, it seems safe to conclude that behavior therapy is as successful for its chosen problems as other therapies seem to be, and in certain specific psychomotor disorders and eneuresis, probably more successful. Research also lays to rest the worry about symptom substitution; treating one symptom does not produce another.

How theoretically sound is behavior therapy? The authors claim to be applying modern learning theory. Though their therapy makes much use of imagery and thought-processes (especially in desensitization), they eschew concepts from perception and cognition. Modern learning theory is not in a state of enough consensus to provide clear explanations for complex phenomena. The authors seem to be using S-R theory as an heuristic analogy rather than as an exact predictive instrument of detailed responses. However, this theoretical inexactness does not detract from the value of the therapeutic techniques that have been created. These techniques can be valuable additions to clinicians' repertories. (The danger is that one becomes rigid in his belief in his preferred scientific mystique and does not recognize that belief itself may have curative powers.) At the worst, it may be true that behavior therapy is more empirical and less rational than the authors wish it to be.

Eysenck and Rachman have presented a very readable and useful first textbook in behavior therapy. The next decade will see many contributions from this approach. Let other clinical persuasions now try to match the behavior therapists' energy and imagination.

University of Oregon NORMAN D. SUNDBERG

Patterns of Redundancy: A Psychological Study. By A. C. STANILAND. Cambridge, Cambridge Univ. Press, 1966. Pp. viii, 216. $8.50.

Readers familiar with Garner's *Uncertainty and Structure as Psychological Concepts* could reasonably expect Staniland's book, appearing four years later, to provide some advances beyond Garner's treatment of essentially the same topics. Readers with this expectation will be disappointed and vexed, in about equal parts. They will find few real gains in data or in the development of useful concepts. Many will find Staniland's clumsy handling of probability-models underlying informational concepts annoying and discouraging. Some confusion about the nature of such models, in fact, underlies much of Staniland's repetitious quarreling with quotations from Garner and the work of others.

But other things must be said about this book. Certainly, few readers will quarrel with the author's intention to explore some existing ambiguities in the use of the term "redundancy," and to relate perceptual performance to the varieties of stimulus-variation that have been described by this single term.

The book provides a description of the computations involved in quantifying the redundancy of some simple, discrete cases of constrained sets. In this, heavy use is made of the graphical properties of constraint; and the discussion is often quite specifically relevant to these properties as depicted in tables and figures. The author describes some familiar cases of constrained stimulus-materials and specifies useful performance-measures, including identification, discrimination, recall, and step-by-step prediction. The measurement of performance in the case of perturbed stimulus-sets (random deletion or substitution of elements) is also considered, and this necessarily involves the description and exploration of good and bad codes (forms of constraint) in logical terms and in terms of the extent to which perceptual performance resists the effects of such perturbation.

The recall task in which the subject views a set of stimuli and then attempts to recall or to produce all the stimuli in the set just seen is given special emphasis. This discussion has considerable intuitive appeal. The task is an easy one if the stimulus-elements in the set to be recalled are all, or most, of those that can be generated by a set of sufficient generation-rules known to the subject. If this is not the case, if the stimuli to be recalled are a small subset drawn unsystematically from a very much larger set, then subjects must rely more on sheer rote memory. The first effect, the role of the subject's knowledge of principles of stimulus-generation or of principles of constraint relating one stimulus in the set to another, is given primary importance; and a serious attempt is made to relate the role of such principles to older notions about memory and familiar beliefs about general perceptual functioning. Much of this discussion is handicapped by unfamiliarity with the work of Garner and his students, subsequent to his book, indicating the role of the specific *form* of constraint in the recall task. Many readers will have little sympathy also for Staniland's extended search for single best measures of "redundancy" that reflect crucial forms of constraint in this context and others. One inconvenience is that the search is confined to unidimensional measures, whereas the complete description of the constraint of a stimulus-set, its *form* and *amount*, necessarily involves more than a single degree of freedom.

Another basic difficulty for the reader is the confusion of *constraint* with *redundancy* in many contexts. The value of the implied distinction can be illustrated in the case of the "metric figures" of Fitts, a case that is elevated to the status of a classical paradox by Staniland. A metric figure is defined as a matrix of sixteen cells arranged in four rows and four columns, each cell being either white or black. The *complete set* of 256 figures is restricted to those in which each column of cells is solidly black to a height of 1, 2, 3, or 4 cells. Two subsets of eight figures each are drawn as samples from the *complete set*. The first, the *random* subset, is formed by unsystematically selecting figures from the *complete set*. The second, *nonrandom*, subset is formed by following a restrictive sampling rule: within a figure, column heights are sampled without replacement. In effect, the *nonrandom* subset is drawn unsystematically from the 24 figures having different solid bar heights in each column. Much metaphysical preoccupation is precluded by reserving the terms "random" and "nonrandom" for the sampling models producing these subsets, rather than applying the terms to the specific subsets produced. Obviously, any subset produced as a *nonrandom* subset can be produced also as a *random* subset, but this need be no real cause for conceptual collapse. Similarly, the application of a measure of *redundancy* to one particular subset produced by either sampling model is bound to confound. Clearly, two kinds of measures are required, one to refer to sampling models producing subsets and another to describe the characteristics of unique subsets produced.

The latter we have. Garner's description of the *total* constraint of a subset, the difference between $\log_2 256$ and $\log_2 8$ in the metric figure model, and the *form* of the constraint, the quantitative assignment of constraint as contingent uncertainties and interaction-uncertainties, is a major contribution. Another measure describing the difference between probability-models underlying the *random* and *nonrandom* sampling processes is needed. Sufficient measures suggest themselves, but whatever the measure used, it must describe the *average* constraint of all subsets that can be generated by a particular sampling model. Considerable conceptual clarification results when the term *redundancy* is reserved, as it has been in earlier works, for such an average measure and the term *constraint* is reserved for characteristics of a particular subset produced by sampling. Any subset generated by the *nonrandom* sampling model, then, whatever its *constraint,* is more *redundant* than any subset generated by the *random* sampling rule. In intuitive terms, any *nonrandom* subset has more in common with all other *nonrandom* subsets thany any *random* subset has in common with all other *random* subsets. Repeated sampling under the random sampling model will generate all possible subsets of the 256 figures taken 8 at a time. Repeated sampling under the nonrandom sampling model will not. It also follows crucially that we may relate the *constraint* of any unique subset of figures to the performance of subjects. But to relate *redundancy* to performance, we must learn how subjects perform with all (or a truly representative sample of all) possible subsets of the same *redundancy.* This is a simple but apparently elusive distinction. And information-theory is relevant to *redundancy* but not to *constraint.*

Readers of Staniland's book will recognize a similar blurring of the distinction between the "goodness" of specific codes and their *redundancy* in the discussion of the perception of perturbed stimulus-subsets. Whether or not the particular assignment of the terms, "constraint" and "redundancy," suggested here satisfies, the distinction described can spare readers much trouble in studying Staniland's book. Regrettably, along with the trouble, much of the apparent point of the book vanishes also; for Staniland would like to be concerned with *redundancy.* In fact, he concerns himself mainly with the effects of the *constraint* of unique subsets, or even with the self-information of a specific subset-element, effects which cannot be generalized in known ways to other subsets of the same kind. This is far short of his apparent real interest, the *redundancy* of the stimulus-world in which we live and our special capacities to respond adequately to it.

This book reminds the reader of a lesson taught by the brief history of information-theory as applied to psychological topics. The lesson has been stated by Frick: information-theory is a formal rather than a substantive theory. It is not a model of behavior; and when it is used in the construction of such models, the application of the theory is appropriate only when the process to be described is intrinsically probabilistic in nature. When the underlying probability-models cannot be described, or are neglected, the application of information analyses is inappropriate and without useful meaning.

University of Illinois

HAROLD W. HAKE

Varieties of Perceptual Learning. By WILLIAM EPSTEIN. New York: McGraw-Hill, 1967. Pp. viii, 323. $9.50.

In his preface, Epstein states that he had three objectives for this volume: (1) to provide a critical review of the perceptual-learning literature; (2) to bring out major unresolved questions in this field, and (3) to arrive at some descriptive generalizations about perceptual learning. These are sensible, though hardly overambitious goals for an author of a book to set himself, but the reader's hopes for enlightenment are quickly dashed when we find him renouncing at the outset the third and presumably most important of these objectives.

This is a direct consequence of Epstein's avoidance of theory, whether of his own making, or of those of others. Perhaps it is unreasonable to have expected the author to have provided us with an original theoretical framework within which to analyze the problems treated in his book, though it surely would have helped. But to systematically skirt around the major theoretical issues in the field (e.g., the controversy between the Gibsons and Postman, the theories of Hebb and Taylor) scarcely seems defensible in a book devoted to the topic of perceptual learning. Even the "general conclusions and proposals" in the brief final chapter amount to no more than some recommendations, largely of a methodological nature, for future research on the major topics dealt with in the book. Sound and cogent though many of the points in this chapter are, they provide little in the way of closure or integration; indeed they would have been far more appropriate as concluding sections to the individual chapters devoted to these topics.

These chapters consist mainly of accounts of research on a diverse and rather arbitrarily selected set of problems in this area, dealt with predominantly in the form of extended presentation and critical discussion of certain key experiments. There is assuredly a place for this sort of book, for those not previously familiar with the field. The book seems indeed to be addressed primarily to advanced undergraduates, or perhaps to beginning graduate students receiving their first brush with the area of perception. For such an audience it may well serve a useful purpose, though most instructors would want to supplement it with a more general treatment, such as Dember's, perhaps, and above all with a good dose of materials directed at the theoretical issues in the field.

Even on this plane, however, there is considerable variation in the adequacy of the various chapters. Epstein is generally at his best in leading the reader through a delimited, relatively well integrated series of studies, such as those on memory color, the Ames trapezoidal window effect, the nature of practice effects in distance judgments and the role of prior exposure of stimuli in the discrimination learning of animals. Where the field becomes more complex, the loose threads more unwieldly, and the competing interpretations more numerous, he tends to lapse into an account of one study after another with little attempt at integration, and to retreat into exhortations for factorial study of additional variables, in formula fashion.

In the process, important issues are apt to become obscured or to be lost sight of altogether. For instance, in the presentation of the material on the perception of size and distance, the critical role of the instructions given the subject is simply acknowledged, without the actual effects of this variable even being mentioned. Similarly, the author describes his own studies on the role of assumed size in the perception of distance, in which positive results were obtained when familiar coins were used, while an attempt to build in assumptions by establishing an artificial correlation between color and size failed to influence perceived distance. Epstein contents himself with a mere restatement of these opposite results, without even a hint at the most likely explanation, that for experimentally established assumptions to become effective, a high degree of overlearning may be required.

While Epstein deliberately disclaims any intent to aim at completeness of coverage of any topic, there are some major omissions which have the effect of leaving the reader with a distorted picture of the state of the evidence. One good case in point is in the emphasis on the role of recency in the perception of ambiguous, Rubin-type pictures, when there is considerable evidence (Botha, Turner and Bevan) of satiation or adaptation effects resulting in a *suppression* of recently perceived figures. It is discouraging, furthermore, to see the foreign-language literature ignored for the most part. The most glaring example is seen in the author's expression of regret concerning the absence of developmental studies in which both size and distance judgments were obtained, when Denis-Prinzhorn, in a monograph cited in a general vein at two earlier points in the same chapter, has reported precisely such data!

All of this left this reviewer pleasantly surprised at the excellence of Epstein's presentation of the work on adaptation to transformed stimulation, to which he dedicates the longest chapter in the book. This is something of a paradox, since this is one of the few topics he deals with to which he has not himself contributed original research. Yet here we find his style loosening up; generous space is given to a presentation of Stratton's and Kohler's pioneering work, in a way which will hardly fail to stimulate the reader's interest in this problem. This leads into an effective and admirably integrated account of the recent research on displaced vision, and of Held's and Harris' views on this subject. On the strength of an impressive marshalling of evidence, the author appears to ally himself squarely with Harris, but refers at the end to Pick and Hay's work on the long-term course of adaptation in the visual and kinesthetic spheres, pointing ultimately to the inadequacy of any view locating adaptation exclusively within either of these systems. Epstein concludes by a trenchant examination of the neglected question of the "why" of adaptation, in the absence of information or reinforcement, and points to a conflict-resolution interpretation that strikes this reviewer as much needed in this area, and which serves to relate this work to other research involving perceptual conflicts, some of which is examined in the preceding chapter of Epstein's book.

It would be nice to be able to close this review on this positive note, but the reviewer would be remiss who failed to mention one further shortcoming of this book, though one for which the author is not to be blamed. This concerns the visual appearance of the text of the book, which is hardly calculated to enhance its readability. Margins so wide they reduce the line of print to 3½ inches, failure to indent paragraphs at the start of a section, headings rendered virtually unreadable by being chopped up into a number of short lines (e.g., The influence of variations of / visual angle on perceived / distance when assumed / size is invariant) — these and other similar defects suggest that McGraw-Hill's designers could profit from an elementary lesson on the principles of perceptual organization!

Clark University JOACHIM F. WOHLWILL

Human Spatial Orientation. By I. P. HOWARD and W. B. TEMPLETON. New York, John Wiley and Sons, Inc. Pp. 533. $13.50.

Nineteen sixty-six was a year for perception books. Forgus and von Fieandt produced general texts: *Perception* and *World of Perception*, respectively. Gibson, Rock, and Howard and Templeton have written three important books for the specialist: *The Senses Considered as Perceptual Systems*, *The Nature of Perceptual Adaptation*, and *Human Spatial Orientation* respectively. Gibson's book is a general theoretical treatise and Rock's is an analysis of a narrowly prescribed problem. The scope of *Human Spatial Orientation* is intermediate. It consists of a comprehensive analysis of a number of problems suggested by the title. The range extends from vestibular sensitivity to geographical orientation, from orientation to gravity to orientation in a weightless state.

The scholarship exhibited in this book is truly remarkable and could well serve as a model to be emulated. The literature reviewed for the various problems includes classical studies from the nineteenth century and a relatively complete survey of the work current at the time of writing (presumably about 1964). A considerable body of German literature is reviewed and frequent references are made to French studies, as well as occasional citation of articles in other languages. The studies are not simply described but rather carefully and critically analyzed. The incisiveness of Howard and Templeton's analysis penetrates the vagueness of some relevant theoretical concepts such as K. U. Smith and W. M. Smith's neurogeometric difference detectors and Werner and Wapner's sensory-tonic theory. Perhaps the authors are too harsh with such approaches—not crediting them with real contributions such as introduction of a most interesting technology in the case of Smith and Smith and the pin-pointing of important problems in sensory interaction in the case of Werner and Wapner.

Howard and Templeton are substantively strongest in constructing taxonomies of some of the persistent problems of orientation. In the introduction the classification of human spatial behavior establishes a vocabulary for analysis of all the subsequent problems considered in the book. A classification of eye movements by function in Chapter 3 sets the stage for an analysis of the relation of eye movements to visual direction. Again in discussing the problem of orientation and shape a simple but often neglected distinction is made between the "effects of shape on perceived orientation" and "the effects of orientation on perceived shape." The latter question is further subdivided into problems concerned with comparisons between shapes whose orientations are manipulated relatively and absolutely and problems concerned with various frames of reference from which to define orientation—the observer, gravity, and the visual framework (p. 295). A final example of a clarifying taxonomy might be the discussion of the meaning of the question, "does the world look upright?" asked of subjects adapting to inverting optical distortions. The authors suggest three different meanings which they term: motor-coordination upright, intersensory upright, and behavioral polarity and they suggest tests to measure the presence of adaptation with respect to each of these meanings.

Some points in the book are disputable. For one thing the authors occasionally answer by assertion what should be considered empirical questions. The most flagrant instance occurs in connection with the adaptation to inverting distortions. The question is raised as to whether a subject can ever "in all senses be said to see the world the right way." The authors suggest one experiment in which distorting devices are worn from birth or for a prolonged period to produce a situation in which "the adapted habits would be the only or strongest standard of normality." They then conclude: "Such experiments have not been done but it is difficult to see how they could do other than prove the correctness of our analysis" (p. 417). Aside from the dubious practice of asserting the results of undone experiments legitimate question might well be raised as to this particular assertion. Certainly there is much evidence in animal research and some from human infants attesting to the innate hookup of some aspects of sensory-motor behavior. Howard and Templeton themselves, in discussing another problem, state: "We suggest that shape recognition involves complex behavior which is largely learned but that visuomotor spatial coordination involves much simpler relationships many of which are probably independent of learning" (p. 343). If some component of visual-motor coordination doesn't depend on experience initially the results of a modification experiment are at the very least not clearly predictable.

Another point that the authors reiterate in connection with adaptation to distortion is that "one cannot talk about localization in a single modality without reference to a motor act or some other modality" (p. 381). It would seem that this assertion neglects the possibility of a "local sign." This logical possibility which is in fact recognized on page 15 of the book would seem to make it possible to talk about localization in a single modality although such a concept probably wouldn't provide the basis for a general explanation of adaptation. It is interesting in connection with local signs that the authors do not discuss localization of points on the body, tasks which might involve tactual-tactual or tactual-proprioceptive localization. Such investigations would hit closest to the classical concept of local sign.

The fact that one can find points of disagreement is of course an impetus to further research. Moreover the authors implicitly with their organization and explicitly by direct suggestion show where gaps in our knowledge exist. It is not at all unlikely that *Human Spatial Orientation* will be a standard item on the shelves of serious students of perception and motor behavior and will become a classic reference book.

University of Minnesota HERBERT L. PICK

Neural Mechanisms of Higher Vertebrate Behavior. By J. S. BERITOFF (BERITASHVILI). (Translated by W. T. LIBERSON). Boston: Little, Brown, and Co., 1965. 384 + xvi. $15.00.

As this is written, Professor Beritoff must be in his 81st year of life and his 57th year of research into the neural mechanisms underlying behavior. His book is important, in fact I believe it to be a "must" for psychologists interested in the learning processes and in the neural apparatus responsible for behavior. It is both an exposition of those theories he has developed over these last 57 years and a summary of many of the experiments on which they are based. They make for exciting reading, and it is unfortunate that most of us in the United States have heard so little about Professor Beritoff before the publication of the present book.

That we haven't heard more of Professor Beritoff is somewhat surprising in itself since the references provided show him to be the only or first author of 82 experimental reports. In passing, it is interesting to observe that he makes use of only two papers by Anokhin and none by Sokolov. Sokolov's recent book *Perception and the Conditioned Reflex,* 1963, provides reference to 8 articles by Anokhin, 12 by Sokolov, and 2 by Beritoff. The differential weighting of referenced materials probably reflects both a geographical and ideologic separation between Anokhin and Sokolov in Moscow and Beritoff in Tiflis.

In over-all perspective the book is an extended essay upon the nature of learning in man and in animals with concrete suggestions for the neural mechanisms associated with molar behavioral changes. Beritoff's views differ in principle from

Pavlovian classical conditioning, since Beritoff believes in three levels of learned behavior. The highest level is that of "conscious," planned behavior which is unique to man. This type of behavior is held to develop after the first few years of childhood. A second level is that of "psychoneural activities," otherwise known as "conditioned images." The third level is that of the classical conditioned reflex. The hierarchy among the systems is such that "level one," that of the conscious, planned behavior is prepotent over all of the other systems. The psychoneural activities of level two, are dominant over the conditioned reflex mechanisms.

Although behavior that is "consciously planned" can be dominant over reflexive or learned behaviors of the other levels, Beritoff develops this aspect of his theory least well. One reason is, of course, that the majority of experiments which provide a basis for his theoretical structure use animals as subjects, and by definition, animals lack the conscious, planned level of activity. Moreover, discourse about long-term, "planned" activities of man seem the most difficult for anyone to place into a meaningful and useful theory. The significance of the theoretical structure of the book is its elaboration of the second level of learned behavior common to both man and other animals, that of image-driven behavior.

In Beritoff's opinion, behavior based on conditioned images differs from conditioned response activities in several ways. He postulates that the neural processes which sustain image-driven behavior are processes which can occur in a single exposure to a situation. This can be interpreted as the all-or-none type of learning which is an attribute of many current mathematical formulations of learning. These same processes are capable of being reproduced by the influence of one component of the situation after considerable periods of time. As a general rule, one might say that Beritoff's laws governing image-driven behavior are similar to those which govern the strength of incentives in most modern learning theories. Despite the similarity in theoretical rules of operation, one should not consider the images of psychoneural behavior to be similar to incentives in other ways. Beritoff believes that images are really projected onto the environment as "constructed" by the neural apparatus of the organism. The constructed environment is based, to considerable extent, upon vestibular information and this point is rather well documented. It also fits with recent work in the United States which demonstrates the great influence of vestibular information in the spatial orientation of animals, generally. Behavior is governed by the conditioned images exactly as it would be if the animal were directly perceiving the object of significance in the environment, because the cells involved in the production of the "mental images" are exactly the same as those responsible for images produced by positive sensory input; namely, the stellate cells of third and fourth layers of "sensory neocortex."

The stellate cell is a short-axoned cortical cell with a dense pericellular reticulum of dendrites and axonal ramifications. They differ from short-axoned pyramidal cells in that they lack apical dendrites and do not have thorny processes coming off their dendrites. These cells play a large role in Beritoff's neural theory. Interesting, if not absolutely compelling, speculations are made concerning the relationships which may exist among the various types of cells, stellate and others, found in the neocortex and related systems, but it is possible to accept the behavioral analyses made using the three levels of learning and not accept the conclusions concerning the functional significance of the cerebral architecture. By the same token, there may be others who will be more excited by the inferences from cellular structure than from the behavioral data. Even if a reader were to hold reservations both about the behavioral and the interpretations of the cerebral cytoarchitecture there is more than enough new and interesting experimental data to make reading the book worthwhile.

The University of Michigan ROBERT L. ISAACSON

The Social Psychology of Organizations. By DANIEL KATZ and ROBERT L. KAHN. New York, John Wiley and Sons, 1966. Pp. ix, 498. $8.50.

This is a welcome contribution to the rapidly developing field of organizational psychology. The major contributions of the book are (1) an original, creative synthesis of the field, and (2) an objective approach that is consistently free of emotional bias. There is reason to be as grateful for the second of these outcomes as for the first.

The authors regard organization as an input-throughput-output system that utilizes both material and human inputs. Organizations develop (1) production subsystems concerned with the work being done; (2) supportive subsystems concerned with procurement, disposal, and institutional relations; (3) maintenance subsystems for tying people into their functional roles; (4) adaptive subsystems concerned with organization change; and (5) managerial systems for the direction, adjudication, and control of the subsystems and activities of the structure.

In addition to the subsystems that perform various functions, organizations develop structural and interactive subsystems based on the differentiation of roles, norms, authority, communication channels, and the like. Norms, rules, authority and sanctions are normal and necessary due to the fact that reliability of role behavior is the requirement intrinsic to human organizations.

Organization is an open system in that it imports members, materials, and values from its environment and exports products and values that affect its environment. The pressures exerted by environmental changes and societal norms require the development of effective systems for gathering and processing information in order to adapt to external conditions. In addition, the internal subsystems generate pressures and contradictory demands that call various feedback mechanisms into operation in order to maintain balance and adaptation. Since differences in authority, values and norms in different segments of the role system operate to block upward and downward communication, information channels tend to become formalized in order to insure the passage of sufficient information for the coördination and control of activities. Four types of dynamics created by the substructures of the organization (the proficiency dynamic, the stability force, the push toward environmental manipulation, and the pressure for change) are often in conflict and must be kept in balance by administrative decision and action. Leadership involves three different functions that are applied to the solution of organizational problems. These are (1) policy making or the guidance and change of structure, (2) piecing out and adapting structure to changing demands; and (3) using existing structure to keep the organization in effective operation.

Every member of the organization is involved in several of its subsystems. Among these are the authority, functional (task specialization), communications, normative, and informal substructures. These structures set bounds to the role that the member is expected to play in the organization. If too closely defined, these bounds may limit the contribution that a member can make to the organization. Some boundary conditions are, however, necessary to permit the individual to contribute toward the purpose of the organization.

The above outline fails to do justice to the richness and complexity of the theoretical system. The structure and dynamic synthesis of the system are in themselves original contributions. The authors call upon their personal observation and experience to add many valuable insights into organization. It is not surprising to find that most of the research of the authors and their colleagues at the University of Michigan is directly relevant to their discussion. Many other scientists whose work could have been used to illuminate various aspects of their system are given scant attention. The contributions of the businessmen and scholars who originated organization theory are ignored.

The objective approach of the authors is deserving of special comment. There is a tendency in the behavioral literature of science to define concepts such as authority and structure in terms that specify the writer's feelings rather than the concept under observation. The authors of this book define authority and structure as normal and inherent aspects of organization. They discuss such concepts in an impartial manner that should serve as a model for other writers in the field.

Ohio State University RALPH M. STOGDILL

Identification and Child Rearing. By ROBERT R. SEARS, LUCY RAU, and RICHARD ALPERT. Stanford, California, Stanford University Press, 1965. Pp. viii, 383. $8.50.

This book is a report of a large-scale research project intended to test a theory of identification derived from a combination of psychoanalytic and Hullian learning theory principles. The principal antecedents of identification were postulated to be the child's dependency motive and parent practices which increase this motive.

Data were gathered from 40 four-year-olds and their parents in a summer nursery school program. Nearly all of the children were from highly educated professional or student families. The resulting homogeniety of the sample may have contributed to the lack of results in the study. Measures included observations of the children, experimental and projective tests, interviews with both parents, observations of the mother with the child, and a questionnaire completed by the mother. From these data, 369 variables describing parent and child behavior were compiled.

Two sets of theoretical predictions were tested. First, dependency, adult role taking, sex typing, conscience development, and aggression were expected to be intercorrelated since they theoretically result from the single mediating process of identification. The results indicated small positive relationships between these variables for girls, but no relationships for boys. Second, parental nurturance, high standards of conduct, use of love-oriented discipline, and verbal presentation of models of behavior were expected to increase behaviors related to identification. The results generally failed to confirm these predictions.

The authors proceeded from formal hypothesis-testing to an inductive analysis by examining all of the parent variables which correlated significantly with a given child behavior. The most interesting of these analyses appeared in the chapter on sex typing. A fairly large number of correlations indicated that, for both boys and girls, parental restrictiveness was associated with femininity while parental permissiveness was associated with masculinity. The authors suggested quite plausibly that parental permissiveness may facilitate the active, exploring, adventurous qualities involved in masculinity.

Although some of the other inductive analyses were provocative, they frequently seemed to go far beyond the data by basing elaborate interpretations on a very few correlations which could easily be the result of chance. In some chapters, this speculative analysis was intermixed with the descriptive presentation of the data in such a way that it was difficult and time-consuming to separate the two. Further, a reformulation of identification theory presented at the end of the book was weaker than the earlier theory and was not consistent with the findings they obtained.

Probably the most important contribution of the book is the disconfirmation of theoretical hypotheses and findings from previous studies (notably R. R. Sears, E. E. Maccoby, H. Levin, *Patterns of Child Rearing,* 1957). Since this study improved on earlier methodology by obtaining independent measures of child and parent behavior and by obtaining several measures of each variable, its findings should supersede earlier results. Although the previous findings had considerable heuristic value, they can no longer be considered conclusions.

On the whole, the book presents some interesting and provocative ideas, but is disappointing because of the relative absence of positive findings. In retrospect, it appears that this failing is partly due to the limited potential inherent in attempts to relate global child rearing variables to global categories of child behavior. Recent developments suggest that a better understanding of children's behavior will result from examination of more specific behavior and from studying immediate stimulus conditions as well as the child's learning history. For example, the study found no correlation between parental reports of the child's behavior and observations in the nursery school. One reason for this discrepancy may be that situational differences produce different behavior. It should be noted that the data were collected in 1958 when the research strategy used looked more promising than it does presently. I suspect, however, that this is the last of the broad studies of child rearing practices which will be attempted for a while.

Cornell University

ALETHA STEIN

An Introduction to Mathematical Learning Theory. By RICHARD C. ATKINSON, GORDON H. BOWER, and EDWARD J. CROTHERS. New York, John Wiley and Sons, 1965. Pp. xii, 429. $9.95.

The stated purpose of this book is to serve as a primer teaching some of the basic mathematical methods that have been used in mathematical learning theory. At this, it must be considered a success. Before its publication, the readership of the growing literature in this field was limited to that small group of psychologists who had somehow managed to provide themselves with the necessary technical background. Atkinson, Bower, and Crothers have given the psychologist with limited mathematical skills a much-needed entry into this literature.

The book is organized around psychological topics, but in such a way that the simpler mathematical models are considered first. The first chapters treat simple models for concept identification and paired-associate learning, with considerable space devoted to insuring an understanding in depth of the kinds of derivations that can be made from these models, with only the most limited mathematical skills presupposed. Later chapters treat more complex learning models under the topics of choice behavior, signal detection and probability learning, avoidance conditioning, social and economic behavior, and stimulus-sampling theory. The book ends with an outstanding chapter on parameter estimation that discusses the general properties of estimators as well as the specific problems involved in estimating parameters in order to assess the goodness-of-fit of various types of models to data.

It is perhaps unfortunate that the authors chose to limit severely their selection of models in order to emphasize the detailed treatment of specialized mathematical techniques. Thus, the models treated are primarily those on which the authors themselves have worked, which in many cases do not give a representative picture of the developments in a particular field. This is particularly true in the chapter on signal detection, which is concerned solely with a stationary high-threshold model that is quite generally considered inadequate by workers in this field. Another unfortunate effect of this emphasis on mathematical detail is the devotion of considerable space to topics such as autocorrelations of errors that many learning theorists would consider the height of trivia. On the other hand, the strict selection of topics gives the reader the opportunity to delve deeply enough into the workings of these models to develop the facility necessary to truly understand the techniques involved in the development of models of this type.

A more severe limitation of the value of the book is due to the authors' failure to devote enough attention to the problem of relating the topics discussed in their book to more of the problems widely considered by non-mathematical psychologists working in the same areas. Thus, the over-all impression given by this book is that mathematical learning theory by nature deals only with its own specialized set of topics. This is a false impression. Mathematical learning theory has a great deal to say, for example, about such things as the basic nature of the learning process and the relation of short-term memory to paired-associate learning. Furthermore, the authors have distorted what mathematical learning theory does have to say by choosing to present only data supportive of the models discussed, with no adequate discussion of their shortcomings.

The general psychologists who reads *An Introduction to Mathematical Learning Theory* for an overview of the important developments and issues in this field will no doubt be disappointed. The book is not a general overview of the area for the slightly interested observer, and it does not attempt to give a representative picture of the field. But the serious student of the application of quantitative models in psychology, who wants to develop the mathematical facility necessary to comprehend and contribute to this field, will find this book extremely useful. Thus, despite its narrowness, the book must be considered a very valuable contribution.

Cornell University HARLEY BORNBACH

The Results of Counseling: Twenty-Five Years Later. By DAVID P. CAMPBELL. Philadelphia, W. B. Saunders Company, 1965. Pp. x, 205. $6.00.

Campbell's central objective was to find out whether a group of men and women who as college freshmen had sought and received short-term counseling (1 to 3 interviews) at the Student Counseling Bureau of the University of Minnesota in the 1930's were, 25 years later, more satisfied with their lot in life and more productive than a group of their peers who had not sought or received such counseling. Each group contained almost 400 subjects, who were willing to give about five hours of their time to complete a 30-page questionary, participate in a one-hour interview, complete the Strong Vocational Interest Blank, and take the current entrance examination of the University of Minnesota. The report is replete with detailed statistics, but the investigator summarizes his principal findings succinctly. Referring to the *counseled* group, he states: "They report slightly less satisfaction with their jobs, their marriages, their social life, and their over-all situation. But they have won more degrees in college, more honors, and have more often been elected to leadership posts. Today they make a few dollars more each year (a statistically insignificant $1,200) and score higher on a gross measure of 'contribution to society'" (pp. 5-6). His general conclusion is that students who received counseling ". . . are slightly better off today than the non-counseled ones" as a result of the counseling they received (p. 25). (The expression "better off" must refer to achievement rather than satisfaction, a value-judgment definition with which not everybody will agree.)

Campbell devotes an entire chapter to an *apologia* intended to support his declaration that in *this* piece of research the lack of a motivated (to seek counseling) control group did not really lead to unwarranted conclusions concerning the efficacy of counseling. The gist of his argument is as follows: because the experimental group was not found to be significantly "better" than the control group (in respect to scholastic aptitude, socio-economic background, and emotional stability), the post-counseling differences in achievement favoring the experimental group should be attributed to the counseling they received rather than to some third factor.

The weakness in Campbell's argument is that he interprets the one persistent difference he did find between the two groups in a way that is decidedly questionable. The counseled participants, not only when they were freshmen but also 25 years later, were found to be somewhat more anxious than the non-counseled subjects. Campbell presents this information as evidence that the experimental group was "worse" than the control group, and states that the achievement results in favor of the counseled subjects are therefore "even more dramatic" (p. 26). He overlooks the research that has been published on the relationship between mild anxiety and achievement; the results of such research suggest that the subjects in the experimental group may have been more strongly motivated to achieve than were their less anxious peers in the control group. The latter group's lack of subjects who were motivated to seek counseling precludes the possibility in this study of drawing a defensible conclusion concerning what effect, if any, the counseling had on subsequent achievement. It *is* clear that the counseling did not reduce the tendency to be anxious of the experimental subjects. One might hypothesize that they would have become anxious to a *crippling* degree if they had not received the counseling, but here again the absence of a satisfactory control group prevents the testing of the hypothesis.

The most useful features of this monograph, in the opinion of the reviewer, are: (1) the description of the "Contribution to Society" rating-scale and its use (pp. 83-92); (2) the sections entitled "Locating 99 Percent 25 Years Later," and "Winning Cooperation (pp. 46-53).

Cornell University A. GORDON NELSON

Interviewing: Its Forms and Functions. By STEPHEN A. RICHARDSON, BARBARA SNELL DOHRENWEND, and DAVID KLEIN. New York, Basic Books, 1965. Pp. viii, 380. $7.50.

For the reader who is looking for a comprehensive treatment of the interview, the title of this book will be misleading. As the reviewer and his co-authors (Walter

V. Bingham, Bruce V. Moore, and John W. Gustad, *How to Interview,* 1959, 7) described it, the interview may have any one or all of three functions: to obtain information, to give information, and to effect changes in behavior. The book under consideration is concerned only with the first of these, obtaining information.

Given this limitation, it is nevertheless a generally well written and thoughtful work. It is organized into four parts. The first deals with the interview as a research instrument and discusses other methods for obtaining information as well. The second part is concerned with respondent-participation and includes consideration of interviewer-effects. The third and largest part has to do with the question-and-answer process and the problems and pitfalls to be found therein. The last part deals with the characteristics of the participants in the interview. A very interesting appendix presents a study by Richardson concerning personality traits (as measured by TAT responses) of social science field workers.

Although fairly frequent references are made to research and opinion reports, the book is primarily a compilation of suggestions and cautions based on substantial experience in a variety of settings. If one had to characterize this book simply, one might best to describe it as a moderately sophisticated field manual for persons interested in using the interview in social research. To profit from it, the reader ought to have some substantial background in interviewing as well as in sociology, anthropology, and psychology. If it is used as a text, it will require considerable supplementation by the instructor.

Even if it is granted that this is more of a manual than a scholarly treatise, the reviewer would still have preferred to see reference made to a good bit more of the rich literature having to do with the interaction between persons. This lacuna is perhaps due to the fact that, despite its title, the book deals with only one kind of interview, whereas the literature referred to above has been produced for the most part by those using the interview in the clinical and counseling setting. One result of this is the description (pp. 314 ff.) of the interviewer-respondent interaction in such terms as "congeniality" and "mutuality."

On the other hand, there is generally skillful use of research materials that are cited. The study conducted by Richardson and Dohrenwend (pp. 183 ff.) and reported in some detail is a highlight of the book. It was concerned with the effects of "leading" questions used both by experienced and inexperienced interviewers and challenges some long-held assumptions and beliefs. It is to be hoped that the authors will continue this line of investigation long enough to tie up the loose ends and perhaps uncover some of the hidden variables.

On balance, this is a useful, readable, and substantively sound book. At a somewhat less technical level, it does much of what the earlier book by Hyman (H. H. Hyman, *Interviewing in Social Research,* 1954) does. The obviously substantial experience of its authors which is woven into the text in properly measured amounts adds to its readability. It will be very worthwhile reading for the journeyman social scientist interested in learning how better to use the interview as a research tool.

The Ohio State University JOHN W. GUSTAD

Shakespeare's Royal Self. By JAMES KIRSCH. New York: G. P. Putnam's Sons, for the C. G. Jung Foundation for Analytical Psychology, 1966. Pp. xi, 422. $7.95.

One of the most over-worked words of the middle part of the twentieth century is certainly the adjective "psychological." It has been used all too frequently by authors and critics whose understanding of its meaning is limited to what they learned in an elementary course, if any. As a result, Kirsch's book will be a pleasant surprise to many readers, for it is a critical evaluation written by a man who studied with Jung and who has also been a practising psychiatrist for forty years.

The "royal self" of the title should not be misinterpreted as applying to Shakespeare's plays about English kings; it is, instead, an analysis of factors in three major tragedies which, Kirsch believes, demonstrate the universality of Shakespeare's mind. Following the preface by Gerhard Adler, Kirsch explains in his introduction that although he has studied and drawn upon most of the Shakespeare canon, he has limited his commentary to the three plays he considers most appropriate for his thesis. These are: "*Hamlet*: A Drama of a Haunted Man"; "*King Lear*: A Play of Redemption"; and "Macbeth's Descent into Hell and Damnation"—presented in the book in that order. The subtitles quoted summarize very well the highly detailed study contained in each section. From Kirsch's point of view, the order of the presentation is logical, running as it does from "heights to depths." From a purely literary point of view, the arrangement is less logical, since it puts in the final, most emphatic position the play which is probably the poorest of all Shakespeare's tragedies because of its faults of structure and inconsistencies.

In spite of that exception, there is no doubt that this book is a carefully planned discussion of extremely complex themes. It has been thoroughly documented with many notes, which make it clear that the author has consulted not only the plays, but a good deal of the scholarship on them. It would, however, have been helpful to the reader if the many references which appear in the notes had been collected into some sort of bibliography at the end of the volume.

The book suggests a number of questions. One is, not who *will* read it, but who *should* read it. Certainly, it will be of interest to psychologists, and advanced graduate students in the humanities will probably find it both interesting and different. But there is one category of reader who should be warned to approach it with caution: the very general reader, or the student whose knowledge of Shakespeare is not yet adequate for reading with critical disinterest. Such a reader will find his mind being made up for him before he can make it up for himself.

Well-written though it is, *Shakespeare's Royal Self* seems to have a major weakness. The author relies for his proof upon minute considerations of word and phrase. This is all very well, but it implies that *Hamlet* and *Lear* and *Macbeth* were written for reading, that they are closet drama, in other words. Yet there is one matter concerning Shakespeare about which there is no doubt: he wrote his plays for production on a public stage, before an audience made up of all classes and degrees of people. He wanted his lines to be understood as they were heard, his characters to be considered as they were presented in the flesh. Four hundred years demonstrate that he succeeded.

It is true that Shakespeare's plays are those rarities, works which can be read almost as profitably as they can be viewed. Dr. Kirsch would have done well to keep both factors in mind.

University of Texas

LUCETTA J. TEAGARDEN

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review).

Abramson, Harold A. (ed.). Introduction by Frank, Fremont-Smith, *The Use of LSD in Psychotherapy and Alcoholism*. New York, Bobbs-Merrill, 1967. Pp. xxv, 697. $17.50.

Appley, Mortimer H., and Trumbull, Richard (eds.), *Psychological Stress: Issues in Research*. New York, Appleton-Century-Crofts, 1967. Pp. xvi, 471. $8.00.

Aronoff, Joel, *Psychological Needs and Cultural Systems. A Case Study*. Princeton, N.J., VanNostrand, 1967. Pp. xiii, 241. $2.95 (paperback).

Babladelis, Georgia, and Adams, Suzanne, *The Shaping of Personality. Text and Readings for a Social-Learning View*. Englewood Cliffs, N.J., Prentice-Hall, 1967. Pp. ix, 522. $5.95 (paperback).

Baldwin, Alfred L., *Theories of Child Development*. New York, Wiley, 1967. Pp. xii, 618. $8.95.

Berenson, Bernard F., and Carkhuff, Robert R. (eds.), *Sources of Gain in Counseling and Psychotherapy. Readings and Commentary*. New York, Holt, Rinehart & Winston, 1967. Pp. ix, 449. $5.95 (paperback).

Berg, Irwin A., *Response Set in Personality Assessment*. Chicago, Ill., Aldine Pub. Co., 1967. Pp. xi, 244. $6.75.

Borger, Robert, and Seaborne, A. E. M., *The Psychology of Learning*. Baltimore, Md., Penguin Books, 1966. Pp. 249. $1.25 (paperback).

Brackbill, Yvonne, and Thompson, George G. (eds.), *Behavior in Infancy and Early Childhood. A Book of Readings*. New York, Free Press, 1967. Pp. xxi, 692. $9.95.

Brehm, Jack W., *A Theory of Psychological Reactance*. New York and London, Academic Press, 1966. Pp. x, 135. $5.75.

Bromley, D. B., *The Psychology of Human Ageing*. Baltimore, Md., Penguin Books, 1966. Pp. 366. $1.45 (paperback).

Cattell, Raymond B. (ed.), *Handbook of Multivariate Experimental Psychology*. Chicago, Ill., Rand, McNally. 1966. Pp. xxxi, 959. $25.00.

Clark, Donald H. (ed.), *The Psychology of Education. Current Issues and Research in Education*. New York, Free Press, 1967. Pp. xiv, 274. $2.95 (paperback).

Cohen, John, *A New Introduction to Psychology*. New York, Hillary House Publishers, 1966. Pp. 220. $5.00.

Conference on the Application of Programmed Instruction to the Hospital Field, *Programmed Instruction and the Hospital. A Report on the Use of Programmed Instruction in Health Care Institutions*. Chicago, Ill., Hospital Research and Educational Trust, 1967. Pp. vii, 155. $6.00.

Conrad, John P., *Crime and its Correction. An International Survey of Attitudes and Practices*. Berkeley and Los Angeles, Calif., Univ. of California Press, 1967 (first published 1965). Pp. viii, 312. $1.95 (paperback).

Cooper, David, *Psychiatry and Anti-Psychiatry*. New York, Barnes & Noble, 1967. Pp. xiii, 128. $4.50.

Duffy, John C., and Litin, Edward M., *The Emotional Health of Physicians*. Springfield, Ill., C. C. Thomas, 1967. Pp. ix, 88. $5.50.

Edgerton, Robert B., *The Cloak of Competence. Stigma in the Lives of the Mentally Retarded*. Berkeley, Calif., Univ. of California Press, 1967. Pp. xviii, 233. $5.75.

Edwards, Allen L., *Statistical Methods*. 2nd ed. New York, Holt, Rinehart & Winston, 1967. Pp. xvi, 462. $8.95.

Fairweather, George W., *Methods for Experimental Social Innovation.* New York, Wiley, 1967. Pp. x, 250. $7.95.

Felix, Robert H., *Mental Illness. Progress and Prospects.* New York, Columbia Univ. Press, 1967. Pp. 110. $4.75.

Ferenczi, Victor, *La Perception de l'Espace Projectif. Étude Réalisée auprès des Travailleurs Africains Analphabètes en France.* Montreal, Paris and Brussels, Didier, 1966. Pp. 146. $4.80.

Foppa, Klaus, *Lernen, Gedächtnis, Verhalten: Ergebnisse und Probleme der Lernpsychologie.* Berlin, Kiepenheuer & Witsch, 1966. Pp. 486. DM 24, 80 (paperback).

Friedman, Arnold P., and Harms, Ernest (eds.), *Headaches in Children.* Springfield, Ill., C. C. Thomas, 1967. Pp. vii, 151. $8.75.

Furst, Sidney S. (ed.), *Psychic Trauma.* New York, Basic Books, 1967. Pp. xv, 252. $6.95.

Games, Paul A., and Klare, George R., *Elementary Statistics. Data Analysis for the Behavioral Sciences.* New York, McGraw-Hill, 1967. Pp. xiv, 564. $8.95.

Gill, Merton M. (ed.), *The Collected Papers of David Rapaport.* New York, Basic Books, 1967. Pp. x, 943. $17.50.

Handel, Gerald (ed.), *The Psychosocial Interior of the Family: A Sourcebook for the Study of Whole Families.* Chicago, Ill., Aldine Pub. Co., 1967. Pp. xi, 560. $10.95.

Harrison, Saul I., and Carek, Donald J., *A Guide to Psychotherapy.* Boston, Mass., Little, Brown & Co., 1966. Pp. xiii, 263. $9.75.

Heckhausen, Heinz (Trans. by Kay F. Butler, Robert C. Birney, and David C. McClelland). *The Anatomy of Achievement Motivation.* New York and London, Academic Press, 1967. Pp. xvii, 215. $5.95.

Heinemann, Robert B., and Bernstein, Marilyn, *The Fourth R. A Return to Learning for Sidetracked Adolescents.* Boston, Mass., Beacon Press, 1967. Pp. viii, 212. $6.00.

Helson, Harry, and Bevan, William (eds.), *Contemporary Approaches to Psychology.* Princeton, N.J., D. VanNostrand, 1967. Pp. xii, 596. $14.50.

Hildum, Donald C. (ed.), *Language and Thought.* Princeton, N.J., VanNostrand, 1967. Pp. vi, 201. $1.95 (paperback).

Hollander, Edwin P., *Principles and Methods of Social Psychology.* New York, Oxford Univ. Press, 1967. Pp. xxiii, 520. $8.00.

Hollander, Edwin P., and Hunt, Raymond G. (eds.), *Current Perspectives in Social Psychology. Readings with Commentary.* 2nd ed. New York, Oxford Univ. Press, 1967. Pp. xii, 685. $5.00 (paperback).

Hook, Sidney (ed.), *John Dewey: Philosopher of Science and Freedom. A Symposium.* New York, Barnes & Noble, 1967 (first published 1950). Pp. vi, 383. $9.50.

Horn, Klaus, *Dressur oder Erziehung: Schlagrituale und ihre gesellschaftliche Funktion.* Frankfurt, Suhrkamp Verlag, 1967. Pp. 127. (paperback).

King, Richard A. (ed.), *Readings for an Introduction to Psychology.* 2nd ed. New York, McGraw-Hill, 1966. Pp. vii, 559. Cloth, $6.50; Paper, $3.95.

Klein, Malcolm W. (ed.), *Juvenile Gangs in Context. Theory, Research, and Action.* Englewood Cliffs, N.J., Prentice-Hall, 1967. Pp. xi, 210. $3.50 (paperback).

Landauer, Thomas K. (ed.), *Readings in Physiological Psychology. The Bodily Basis of Behavior.* New York, McGraw-Hill, 1967. Pp. xii, 584. Cloth, $7.50; Paper, $5.50.

Lang, Alfred (ed.), *Rorschach-Bibliographie 1921-1964.* Bern und Stuttgart, Verlag Hans Huber, 1966. Pp. 191. Fr./DM 36.—

Leichter, Hope Jensen, and Mitchell, William E., *Kinship and Casework.* Russell Sage Foundation, 1967. Pp. xxi, 343. $7.50.

Lemert, Edwin M., *Human Deviance, Social Problems, and Social Control.* Englewood Cliffs, N.J., Prentice-Hall, 1967. Pp. viii, 211. $3.50 (paperback).

Lenneberg, Eric H., *Biological Foundations of Language.* New York, Wiley, 1967. Pp. xvi, 489. $14.95.

LINDGREN, HENRY CLAY, *Educational Psychology in the Classroom.* 3rd ed. New York, Wiley, 1967. Pp. xx, 686. $8.50.

LIEBERMAN, PHILIP, *Intonation, Perception, and Language.* Cambridge, Mass., M.I.T. Press, 1967. Pp. xiii, 210. $7.50.

LOWENSTEIN, OTTO, *The Senses.* Baltimore, Md., Penguin Books, 1966. Pp. 216. $1.25 (paperback).

LUCHTERHAND, ELMER and SYDIAHA, DANIEL, *Choice in Human Affairs. An Application to Aging-Accident-Illness Problems.* New Haven, Conn., College & University Press Services, 1966. Pp. 176. Cloth, $5.00; Paper, $1.95.

LURIA, A. R. (Trans. and ed. by W. Horsley Gantt), *The Nature of Human Conflicts.* New York, Washington Square Press, 1967 (first published 1932). Pp. xvii, 431. $1.45 (paperback).

MAHER, BRENDAN A. (ed.), *Progress in Experimental Personality Research.* Vol. 3. New York and London, Academic Press, 1966. Pp. xii, 319. $9.50.

MALPASS, LESLIE (ed.), *Social Behavior. A Program of Self-Instruction.* New York, McGraw-Hill, 1967. Pp. xix, 489. Cloth, $7.95; Paper, $5.50.

MCCANDLESS, BOYD R., *Children. Behavior and Development.* 2nd ed. New York, Holt, Rinehart & Winston, 1967. Pp. xiv, 671. $8.95.

MILLER, GEORGE A., *The Psychology of Communication. Seven Essays.* New York, Basic Books, 1967. Pp. vii, 197. $4.95.

MILLER, HARRY L. (ed.), *Education for the Disadvantaged. Current Issues and Research in Education.* New York, Free Press, 1967. Pp. xii, 290. $2.95 (paperback).

MILLON, THEODORE (ed.), *Theories of Psychopathology. Essays and Critiques.* Philadelphia, Pa. and London, W. B. Saunders, 1967. Pp. xv, 444. Cloth, $7.50: Paper, $4.25.

MOUSTAKAS, CLARK, *Creativity and Conformity.* Princeton, N.J., VanNostrand, 1967. Pp. xv, $1.95 (paperback).

NEWCOMB, THEODORE M., KOENIG, KATHRYN E., FLACKS, RICHARD, and WARWICK, DONALD P., *Persistence and Change. Bennington College and Its Students After 25 Years.* New York, Wiley, 1967. Pp. viii, 292. $9.95.

OPLER, MARVIN K., *Culture and Social Psychiatry.* New York, Atherton Press, 1967. Pp. xxvi, 434. $8.95.

PASAMANICK, BENJAMIN, SCARPITTI, FRANK R., and DINITZ, SIMON, *Schizophrenics in the Community.* New York, Appleton-Century-Crofts, 1967. Pp. xv, 448. $8.00.

PIAGET, JEAN, and INHELDER, BÄRBEL, *The Child's Conception of Space.* New York, Norton, 1967 (first published 1948). Pp. xii, 490. $2.95 (paperback).

PIKE, KENNETH L., *Language in Relation to a Unified Theory of the Structure of Human Behavior.* The Hague, Mouton & Co., 1967. Pp. 762. 72.- Guilders.

POLLIO, HOWARD R., *The Structural Basis of Word Association Behavior.* The Hague, Mouton & Co., 1966. Pp. 95. 15.- Guilders (paperback).

POPHAM, W. JAMES, *Educational Statistics. Use and Interpretation.* New York, Harper & Row, 1967. Pp. x, 429. $8.95.

PUTNAM, CARLETON, *Race and Reality. A Search for Solutions.* Washington, D.C., Public Affairs Press, 1967. Pp. 192. $4.50.

QUIRK, RANDOLPH, and SVARTVIK, JAN, *Investigating Linguistic Acceptability.* The Hague, Mouton & Co., 1966. Pp. 117. 14.- Guilders (paperback).

RABKIN, LESLIE Y., and CARR, JOHN E., *Sourcebook in Abnormal Psychology.* Boston, Mass., Houghton Mifflin, 1967. Pp. vii, 519. $4.95 (paperback).

RAPPEPORT, JONAS R. (ed.), *The Clinical Evaluation of the Dangerousness of the Mentally Ill.* Springfield, Ill., C. C. Thomas, 1967. Pp. xviii, 118. $6.50.

RHEINGOLD, JOSEPH C., *The Mother, Anxiety, and Death. The Catastrophic Death Complex.* Boston, Mass., Little, Brown & Co., 1967. Pp. xi, 271. $9.50.

ROBERTS, JOAN I. (ed.), *School Children in the Urban Slum. Readings in Social Science Research.* New York, Free Press, 1967. Pp. xiii, 626. $7.50.

ROGERS, CARL R. (ed.), *The Therapeutic Relationship and its Impact. A Study of Psychotherapy with Schzophrenics.* Madison and Milwaukee, Wisc., and London, Univ. of Wisconsin Press, 1967. Pp. xix, 625. $12.50.

ROGERS, RAYMOND (Intro. by S. I. HAYAKAWA). *Coming into Existence. The Struggle to Become an Individual.* Cleveland, O., World Publ. Co., 1967. Pp. vii, 128. $4.95.

RUITENBEEK, HENDRIK M. (ed.), *Psychotherapy of Perversions. Studies in Depth of the Whole Literature of Perversions with Representative Case Histories.* New York, Citadel Press, 1967. Pp. 474. $7.95.

RUSSELL, ROGER W. (ed.), *Frontiers in Physiological Psychology.* New York and London, Academic Press, 1966. Pp. xv, 261. $8.50.

SANDSTROM, C. I., *The Psychology of Childhood and Adolescence.* New York, Barnes & Noble, 1967 (first published 1961). Pp. 259. $6.50.

SINCLAIR-DE ZWART, H., *Acquisition du Language et Développement de la Pensée. Sous-systèmes Linguistiques et Opérations Concrètes.* Paris, Dunod, 1967. Pp. 168. 23F (paperback).

SLAMECKA, NORMAN J. (ed.), *Human Learning and Memory. Selected Readings.* New York, Oxford Univ. Press, 1967. Pp. x, 543. $5.75 (paperback).

STAGNER, ROSS (ed.), *The Dimensions of Human Conflict. The Franklin Memorial Lectures, Vol. XV.* Detroit, Mich., Wayne State Univ. Press, 1967. Pp. 194. $4.95.

STEKEL, WILHELM (Trans. and ed. by EMIL A. GUTHEIL), *Compulsion and Doubt.* New York, Washington Square Press, 1967 (first published 1949). Pp. xii, 608. $1.45 (paperback).

STEVENSON, HAROLD W., HESS, ECKHARD H., and RHEINGOLD, HARRIET L. (eds.), *Early Behavior: Comparative and Developmental Approaches.* New York, Wiley, 1967. Pp. ix, 303. $9.75.

STOLZ, LOIS MEEK, *Influences on Parent Behavior.* Stanford, Calif., Stanford Univ. Press, 1967. Pp. viii, 355. $8.95.

SWANSON, GUY E., *Religion and Regime. A Sociological Account of the Reformation.* Ann Arbor, Mich., Univ. of Michigan Press, 1967. Pp. x, 295. $7.50.

TALLENT, NORMAN, *Psychological Perspectives on the Person.* Princeton, N.J., VanNostrand, 1967. Pp. x, 293. $2.95 (paperback).

THORPE, EARL E., *Eros and Freedom in Southern Life and Thought.* Durham, N.C., Seeman Printery, 1967. Pp. xii, 210. $7.75.

USDIN, GENE L. (ed.), *Adolescence: Care and Counseling.* Philadelphia, Pa., Lippincott, 1967. Pp. 238. $7.00.

VENABLES, P. H., and MARTIN, IRENE (eds.), *Manual of Psycho-Physiological Methods.* New York, Wiley, 1967. Pp. ix, 557. $12.50.

WALLIN, J. E. WALLACE, *The Early Development of Special Educational Facilities for Handicapped Children in the Delaware Public Schools.* Wilmington, Dela., J. E. Wallace Wallin, 1966. Pp. 30. 50c (paperback).

WATSON, JOHN B. (Introduction by R. J. HERRNSTEIN), *Behavior. An Introduction to Comparative Psychology.* New York, Holt, Reinhart & Winston, 1967 (first published 1914). Pp. xxxvi, 439. $7.95.

WEINTRAUB, DANIEL J., and WALKER, EDWARD L., *Perception.* Belmont, Calif., Brooks/Cole Pub. Co., 1966. Pp. viii, 103. Trade, $2.60; Text, $1.95 (paperback).

WHALEN, RICHARD E. (ed.), *Hormones and Behavior.* Princeton, N.J., VanNostrand, 1967. Pp. vi, 266. $2.95 (paperback).

WRIGHT, HERBERT F., *Recording and Analyzing Child Behavior. With Ecological Data from an American Town.* New York, Harper & Row, 1967. Pp. viii, 291. $4.50 (paperback).

WYSZECKI, GUNTER, and STILES, W. S., *Color Science: Concepts and Methods, Quantitative Data and Formulas.* New York, Wiley, 1967. Pp. xiv, 628. $27.50.

THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. STANLEY HALL

VOL. LXXX DECEMBER, 1967 No. 4

AN EDITORIAL ANNOUNCEMENT

With this number, I turn over my long 'stewardship' of THE AMERICAN JOURNAL OF PSYCHOLOGY to my Alma Mater, The University of Illinois. I do this now because, through the accidents of publication and of birth, the JOURNAL and I saw the light of day eighty years ago, in the fall of 1887. Since then the JOURNAL has been published from four universities: the first two volumes were from the Johns Hopkins University, Baltimore, Maryland; the next 29 from Clark University, Worcester, Massachusetts; the next 30 from Cornell University, Ithaca, New York; and the last 19 from the University of Texas, Austin, Texas. Counting the five years I served as business editor in 1921-1925, I have seen 49 volumes and two General Indexes (Vols. 1-30) and (31-50) and a third (Vols. 51-75) soon to be released, through the presses. It is time, therefore, that the JOURNAL be given a permanent home—that its wandering about the country be ended; and also that it be placed in enduring hands by which it will be published in perpetuity in the interest of experimental psychology for which it was counted and to which it has since been devoted.

Three times in the JOURNAL'S history it has been edited for short periods by individuals—by G. Stanley Hall, its founder, Vols. 1-6, 1887-1895; E. B. Titchener, Vols. 32-35, 1921-1925; and by me, Vols. 64-67, 1951-1954; and three times for longer periods by boards of co-editors—first, Hall, Titchener, E. C. Sanford, and J. W. Baird from 1895-1920; second, M. F. Washburn, Madison Bentley, E. G. Boring, and me from 1926-1950; and third, M. E. Bitterman, E. B. Newman, Leo Postman, and me from 1955-1967. It is my considered opinion, having acted in both capacities, that the JOURNAL was served best and psychology has benefited most when its editing was a joint affair.

In arranging this transfer of the JOURNAL, I have asked for and have received no assurances or guarantees. No one is wise enough to make decisions regarding the future, particularly when 'perpetuity' is involved. I hope, however, that the management of the JOURNAL will be placed in the hands of a board of experimental psychologists who are nationally selected and appointed for indeterminate periods.

KARL M. DALLENBACH

STUDIES IN CREATIVITY: INDIVIDUAL DIFFERENCES IN THE STORING AND UTILIZATION OF INFORMATION

By N. R. F. MAIER, MARA JULIUS, and JAMES A. THURBER,
University of Michigan

Creativity is often defined as a reorganization of experience, but the evidence for such reorganization is based upon an examination of the end product. Thus the product of creative thinking rather than the process or the mechanism involved is examined. Discussions of learning likewise make reference to the product achieved (memories), but in this instance the mechanisms involved are also described: *i.e.* formation of associative bonds. What are the possible mechanisms of creativity?

Mednick approaches creativity by regarding it as the product of remote associations.[1] Evidence cited takes the form of the richness of associative bonds, which is measured by the number of associations a given memory item will elicit. In order, however, to account for the creative product, he in addition postulates: (1) a creative hierarchy, (2) lack of stereotyped associations, (3) personality styles, and (4) selection of creative combination sets. The last three factors are difficult to translate into learning mechanisms and raise the question of whether creativity can be satisfactorily explained on an associative basis. Wertheimer stresses the reorganization which is achieved through insight.[2] He does mention a mechanism, termed "closure," which refers to stresses in the psychological field caused by the problem's gap or incongruity. Maier postulated a mechanism, "direction in thinking," to account for the selection of parts of past experience and their combination into new meaningful patterns.[3] The vagueness of these mechanisms indicates the need for basic research in the way information is stored and used by the organism.

Most studies of creativity avoid the question of a mechanism or process and are concerned with individual differences and personality variables. Interestingly enough, this lack of concern with finding a mechanism is not true of the literature on concept-formation. In this case, there is fair agreement that the process of concept-formation is one in which the common element in a group of particulars receives more reinforcement than the other elements and so is abstracted or dis-

---

* Received for publication February 7, 1967. The research reported here was supported by USPHS Grant No. MH-02704, United States Public Health Service. We are indebted to Dr. John E. Milholland, Dr. Ronald J. Burke, Junie Janzen, and Melba Colgrove for helpful suggestions.

[1] S. A. Mednick, The associative basis of the creative process, *Psychol. Rev.*, 69, 1962, 220-232.

[2] Max Wertheimer, *Productive Thinking*, 1959, 234-259.

[3] N. R. F. Maier, Reasoning and learning, *Psychol. Rev.*, 21, 1931, 332-346.

tilled out. Solutions to problems reached by trial and error and by means of the transfer of training also suggest rather specific underlying mechanisms. Since the above-mentioned processes can be reduced to learning concepts (associative mechanisms), they are accepted with less critical scrutiny. Why the neglect in designating a mechanism in the case of creative processes?

Two possibilities suggest themselves: One school of thought assumes that the creative mechanism is basically the same as for learning and therefore represents a particular aspect of the learning function; the other school of thought assumes that creativity is a higher mental process than learning and so becomes a gift that some have and others lack. This latter approach tends to associate creativity with gifted individuals, without the process being regarded as something that needs to be explained.

Nevertheless, points of view as different as those of Mednick and Wertheimer agree that the creative product involves the formation of new combinations (Mednick speaks of mediating connecting links, while Wertheimer speaks of insightful new relations). If, indeed, new combinations of old products do occur, it would seem that a direct test of this could be made without inferring it from the analysis of creative products. The testing of the existence of an underlying mechanism for generating new products therefore seems to be an appropriate place to begin a study of the reorganization of experience.

When one examines the concept of a reorganization of experience, it appears at first sight to be in contradiction with the concept that experiences are organized according to the way they are learned. If experiences are organized according to the strengths of the various bonds, then new and unusual combinations should not occur. It is conceivable, however, that strong bonds might be superseded if there is a summation of bonds. Thus Mednick might argue that three remote associative bonds with the same element might have sufficient strength to recall this element and to replace a more immediate association. He does not, however, choose to use this possibility, using instead personality variables and trial and error selection concepts. Another possible way in which this apparent contradiction could be reconciled is through the discovery of mechanisms in addition to that of association formation.

Does reorganization actually occur and do new combinations sometimes replace older and better-learned ones? If such be the case, then two assumptions suggest themselves. (1) Associative bonds may be fragmented so as to make them available for new combinations.[4] (Fragmenting is not to be confused with forgetting or extinction, which are concepts related to associative mechanisms.) (2) Forces other than reinforcement operate to bring two weakly- or non-associated elements together under appropriate conditions. Both of these assumptions conflict with associative mechanisms, and if such additional mechanisms are to be considered, some well-grounded evidence must be forthcoming. First of all, if organization through learning and reorganization through problem-solving are different, individual differences in the relative strengths of these unique processes should become apparent. If, on the other hand, this line of thinking is purely a verbal trick, then these different ways of handling a problem should cause individuals to differ from each other in degree rather than in kind. More specifically, we are raising the question of whether some persons are relatively more inclined to use elements in thinking that have been associated in past experience, whereas others are relatively more inclined to neglect associative bonds or to make new combinations of old elements. If these tendencies are qualitatively different, then measures of them should be either uncorrelated or negatively correlated.

---

[4] Maier, Selector-integrator mechanisms in behavior, in Bernard Kaplan and Seymour Wapner (eds.), *Perspectives in Psychological Theory,* 1960, 135-172.

The purpose of this investigation is to determine whether the process of storing and using previously learned information differs in individuals in a qualitative manner. In other words, if pairs of words are learned by a group of *S*s, will these pairs tend to reveal themselves in stories written by the *S*s or will new pairs be generated by fragmenting and making new pairs out of the learned pairs?

METHOD

Procedure. The training situation required the learning of 12 pairs of words. These pairs were presented on cards to groups of approximately 12 *S*s each. Each pair was exposed for 2 sec. The order of presentation of the word-pairs was varied in a random order for each of the presentations. A total of 12 repetitions were given.

Stimulus-words. The list of stimulus-words is shown below in the sequence in which they appeared on the first presentation.

1. box cracker
2. damp weather
3. steal prison
4. diamonds warden
5. cold cheese
6. bar punishment
7. crime people
8. water warm
9. hot soup
10. dry basement
11. bread milk
12. wicked drunkard

A memory-test (exposing the first member of each pair) was made after the 4th, 8th, and 12th repetitions. The *S*s wrote their responses on blank sheets of paper. More than two-thirds of the *S*s made perfect scores after 8 repetitions so that 12 repetitions were more than adequate for learning the words.

At the end of the learning period, the *S*s were asked to write a story about a prisoner, using as many of the words from the list as possible. The words could be used in pairs, singly, or in any combination that made sense. The *S*s were informed that the stories would be graded on originality and the number of experimental words used. They were not allowed to reproduce the list of words in preparation for the story assignment. A period of 15 min. was allowed for writing the story, but most of the *S*s finished sooner.

After completing the first story, the *S*s were asked to write a second story on a topic of their own choice. This story would be graded in the same way as the first story. The above-mentioned time-limit was also imposed for this story.

The pairs of experimental words are of three types: (a) pairs that are Neutral in that they cannot easily be used in proximity (Pairs 3, 4, 6, 7); (b) pairs that are Congruous in that they readily fit together in the story assigned (Pairs 1, 2, 5, 12); and (c) Pairs that are Incongruous with the assigned story topic (Pairs 8, 9, 10, 11).

Subjects. Undergraduate students from the pool of *S*s (students electing introductory psychology) participated in this experiment. Because men and women *S*s perform differently in a variety of ways, the data from the sexes have been treated separately although data were collected from mixed groups.[5] A total of 67 men and 96 women participated in these experiments.

Scoring of stories. The stories were scored as follows: (1) upon length (total number of words written); (2) upon number of experimental words used; (3) upon number of learned pairs used (Old-Pair score); (4) upon number of pairs fragmented and a single member used (Single-Word score); (5) upon number of new pairs created from the list of words (New-Pair score); and (6) upon the difference between Old-Pair and New-Pair scores (Difference-Score). The above scores on the various ways in which the experimental words were used do not include repeated usage in the same score category. Thus *S* could not magnify a particular score by the mere duplication of a particular word or combination of words.

Memory-test scores were also collected. These were based on the total number of pair members correctly recalled (first member of pair exposed) on the memory-tests given after the 4th, 8th, and 12th repetitions.

The stories written in our experiment permitted creativity of the type displayed in the stories used by Getzels and Jackson.[6] We therefore adopted their six categories for rating the creativity of our stories. Both stories were scored on a 5-point scale on each category. Two graduate students independently rated the stories. Since the correlation between raters was 0.802 (Pearson's *r*, significant at 0.01 level), the two ratings were combined so that the maximal score per story was 60 rather than 30 points.

Stories were also rated for dissimilarity. The theme for the story was specified only for the first story ("Prisoner"); for the second story it was a free choice. The dissimilarity between the stories was scored on a 5-point scale separately on two dimensions: (1) dissimilarity of topics; (2) dissimilarity of direction of the plot. The correlation between raters was 0.851 for dissimilarity of topic and 0.639 for dissimilarity of direction. Since the agreement of the raters was high, the ratings were combined, making the maximal score of dissimilarity 10 for each dimension.

External scores. In addition to the above-mentioned experimental measures, we obtained scores from the following tests and scales.

(1) Guilford Gestalt Transformation Test.[7]

(2) Scholastic Aptitude Test: Verbal and Mathematical subtests.[8]

(3) Opinion, Attitude, and Interest Survey: Achiever Personality Scale (measure of personality attributes associated with academic success); Intellectual Quality Scale (measure of personality attributes associated with intellectual orientation); and Creative Personality Scale (measure of personality attributes associated with creative behavior).[9]

---

[5] Data on sex differences in problem-solving can be found in L. R. Hoffman and N. R. F. Maier, Social factors influencing problem solving in women, *J. Pers. Soc. Psychol.,* 4, 1966, 382-390.

[6] J. W. Getzels and P. W. Jackson, *Creativity and Intelligence,* 1963, 37-43.

[7] J. P. Guilford and P. R. Christensen, A factor-analytic study of verbal fluency, *Rep. Psychol. Lab.,* University of Southern California, 1957, (No. 18), 51.

[8] College Entrance Examination Board, *Twenty-Sixth Annual Report of the Secretary: The Examinations of June 1926,* 1926, 3-9; College Entrance Examination Board, *A Description of the College Board Scholastic Aptitude Test,* 1956, 1-53.

[9] Benno Fricke, *Opinion, Attitude, and Interest Survey Handbook: A Guide to Personality and Interest Measurement,* 1963, 3, 75-87.

RESULTS

Significance of types of pairs. In this analysis, the *Ss'* use of a word or pair was counted only once for a given story. Although we also obtained scores in which repeated usages of the same word or pair were counted, the two methods yielded comparable results.

It will be recalled that the word-pairs were classified as Neutral, Congruous, and Incongruous. The words in each of these classifications could be used as Old-Pairs (by keeping a given pair intact), as Single-Words (by fragmenting the pair), or as New-Pairs (by combining elements of different pairs). For this analysis, we are concerned with determining whether these three classifications of pairs were related to the way the words were used. In scoring the usage of Old-Pairs we merely counted the frequency with which each type of pair was kept intact. Single-Word scores were derived by counting the frequency with which either member of the pair was used. Since there are twice as many items to score (24 words vs. 12 pairs), this measure is necessarily larger. In scoring for New-Pairs, any pair formed could have members from one or two of the classifications. We scored these by counting the times each member of the new pair had its origin in a given classification. Since we are concerned with the proportion of times each type of word usage was associated with each classification, the absolute magnitude of the possible scores have no effect on these measures.

TABLE I

USE OF THREE TYPES OF WORD-PAIRS BY THE MEN
($N=67$)

Pair classification	Old-Pairs	Single-Words	New-Pairs
Story 1:			
Neutral Pairs (Nos. 3, 4, 6, 7)	18 (6.7%)*	367 (55.4%)*	156 (33.3%)
Congruous Pairs (Nos. 1, 2, 5, 12)	121 (45.3%)*	152 (23.0%)*	153 (32.7%)
Incongruous Pairs (Nos. 8, 9, 10, 11)	128 (48.0%)*	143 (21.6%)*	159 (34.0%)
Total	267† (100%)	662 (100%)	468 (100%)
Story 2:			
Neutral Pairs	10 (5.0%)‡	339 (51.5%)‡	125 (27.3%)
Congruous Pairs	98 (49.2%)‡	163 (24.8%)‡	158 (34.5%)
Incongruous Pairs	91 (45.7%)‡	156 (23.7%)‡	175 (38.2%)
Total	199† (100%)	658 (100%)	458 (100%)

* Difference in percentage for similarly superscripted entries significant at 0.001 level.

† Decline in Old-Pairs from Story 1 to Story 2 significant at 0.01 level.

‡ Difference in percentage for similarly superscripted entries significant at 0.001 level.

The top half of Table I shows the frequency with which the men used each type of pair (or one of its members) in Story 1, *i.e.* Old-Pair, a Single-

Word, and a New-Pair. As might be expected, the Neutral pairs were used as Old-Pairs less frequently (6.7%) and as Single-Words more frequently (55.4%) than either the Congruous or Incongruous pairs. The nouns (in Neutral pairs) are more difficult to use together even though scoring did permit the use of a connecting word between them. Thus, *bread and milk* and *people's crime* were counted as Old-Pairs. Nevertheless, the usage of Neutral pairs as Old-Pairs made up only 6.7% of the total number of Old-Pairs. Neutral pairs accounted for more than half of the Single-Word scores. The differences between the Neutral and Congruous pairs and between the Neutral and Incongruous pairs are highly significant. With respect to the New-Pair score, the three classifications of words showed no meaningful difference.

The pairs classified as Congruous and Incongruous were used in strikingly similar manners. It was hypothesized that Incongruous pairs would be fragmented because they were incongruous with the assigned subject of the first story. Although this incongruity was experienced by many *S*s, it was solved by such simple methods as "the prisoner did not have a *dry basement* cell"; "He did not get *hot soup* for his meals." Thus the tendency of some *S*s to retain Old-Pairs was indicated by this negation. In contrast, other *S*s made New-Pairs out of Old-Pairs even when they were congruous with the topic of the story.

In Story 2 the topic was not assigned, hence the congruity of word-pairs was not controlled by the *E*s. The lower half of Table I shows the results for the men. Comparison of the stories reveals a striking similarity in the usage of the three types of word-pairs. Neutral pairs again appeared more often as Single-Words, whereas Congruous and Incongruous pairs received similar treatment. There is no indication that the topic of the story influenced the use of the memorized pairs, *i.e.* whether used together, fragmented, or reorganized. Only one significant difference emerges from a comparison of the stories. Old-Pairs are used less frequently in the second story than in the first, ($p < 0.01$), while Single-Words and New-Pairs show only a slight and insignificant reduction.

The data from the women are shown in Table II and reveal similar results. The same significant differences in word usage for the different types of pairs are present. Again a decline in the use of Old-Pairs between Stories 1 and 2 is found, supporting the point that a trend occurred opposite to what one might expect due to task repetition.

The influence of Neutral pairs on reorganization. We have shown that Neutral pairs contributed more than their share to Single-Word scores and less than their share to Old-Pair scores. This means that word usage may not only be a function of individual differences in preference, but also may be influenced by our selection of word pairs. Since we are concerned with differences in the ways the *S*s use the words, we scored each individual's stories in two ways: first, with all pairs included; second, with Neutral pairs excluded.

TABLE II

USE OF THREE TYPES OF WORD-PAIRS BY THE WOMEN
($N = 96$)

Pair classification	Old-Pairs	Single-Words	New-Pairs
Story 1:			
Neutral Pairs (Nos. 3, 4, 6, 7)	16 (4.6%)*	554 (53.4%)*	263 (37.1%)
Congruous Pairs (Nos. 1, 2, 5, 12)	155 (44.8%)*	269 (26.0%)*	236 (38.2%)
Incongruous Pairs (Nos. 8, 9, 10, 11)	175 (50.6%)*	214 (20.6%)*	211 (29.7%)
Total	346† (100%)	1037 (100%)	710 (100%)
Story 2:			
Neutral Pairs	19 (6.8%)‡	527 (52.1%)‡	196 (28.7%)
Congruous Pairs	133 (47.3%)‡	260 (25.7%)‡	236 (34.6%)
Incongruous Pairs	129 (45.9%)‡	225 (22.2%)‡	250 (36.7%)
Total	281† (100%)	1012 (100%)	682 (100%)

* Differences in percentage for similarly superscripted entries significant at 0.001 level.

† Decline in Old-Pairs from Story 1 to Story 2 significant at 0.01 level.

‡ Difference in percentage for similarly superscripted entries significant at 0.001 level.

TABLE III

CORRELATIONS (PEARSON) BETWEEN MEASURES FOR 12 AND 8 PAIRS

Classification	Men		Women	
	Story 1	Story 2	Story 1	Story 2
Old-Pairs	.968	.980	.986	.971
Single-Words	.896	.905	.894	.899
New-Pairs	.818	.890	.908	.907

Note: All correlations significant beyond the 0.01 level.

To determine the extent to which the Neutral pairs influenced the differences in the usage of words, individual scores based upon 12 pairs were correlated with individual scores based upon 8 pairs (4 Neutral pairs excluded). Table III shows all correlation coefficients to be above 0.8, clearly indicating that the individuals, rather than the inclusion of Neutral pairs, determined the relative magnitude of the individual scores. Thus the scores based on 8 pairs and on 12 pairs revealed basically the same pattern of individual differences in both stories and for both sexes. It appears, therefore, that insofar as our selection of words influenced their usage, the influence on our *S*s was very similar and hence did not determine the way individuals reacted to the learned pairs.

The correlation between the frequencies with which Old-Pairs, Single-Words, and New-Pairs were used in the two stories indicate the extent to which these measures are determined by the individual characteristics, rather than by story topic and random factors. These correlations are shown in Table IV. The correlations are all significant and positive, and demonstrate that the tendencies to use Old-Pairs, Single-Words, and New-Pairs are to a considerable extent a function of individual characteristics. Again, measures based on 12 pairs and 8 pairs show similar relationships, particularly with regard to Old-Pairs and New-Pairs, which are our most crucial measures.

TABLE IV

CORRELATIONS BETWEEN PAIR USAGE IN STORY 1 AND STORY 2

Correlations	Story 1 vs. Story 2	
	Men	Women
Story-Length	.702	.705
Experimental-Words (12 pair)	.739	.517
Experimental-Words (8 pair)	.659	.414
Old-Pairs (12 pair)	.645	.585
Old-Pairs (8 pair)	.639	.596
Single-Words (12 pair)	.512	.373
Single-Words (8 pair)	.403	.439
New-Pairs (12 pair)	.683	.396
New-Pairs (8 pair)	.672	.411
Difference-Score (OP minus NP) (8 pair)	.717	.523

Note: All correlations are significant beyond 0.01 level.

Similarity of word usage in the two stories. If individuals tend to use a similar proportion of Old-Pairs, Single-Words, and New-Pairs when writing the two stories, to what extent do the pairs of words elicit similar reactions in our *S*s with regard to whether the pair-members are used together, require fragmenting, or become parts of new pairs? To determine this relationship, each of the 24 words was scored in terms of how often it appeared as a member of its Old-Pair, as a Single-Word, and as a member of a New-Pair. The rank order correlations for the frequency of these three usages in the stories are shown in Table V.

All correlations are significant at or beyond the 0.05 level of confidence, indicating that the nature of the pairs of words influenced to a considerable degree how *S*s would use them in the two stories (*i.e.* those words used as Old-Pairs in Story 1 also tended to be used as Old-Pairs in Story 2, and so on). It should be noted, however, that the correlation in most instances is lower for the 8-pair comparison than for the 12-pair. Thus, as already indicated, the nature of the Neutral pairs partly influences their usage, but in addition each of the other pairs encourages a particular word usage which is consistent for the two stories.

TABLE V

CORRELATION OF WORDS USED AS OLD-PAIRS, SINGLE-WORDS, AND NEW-PAIRS IN STORIES 1 AND 2

Story 1 vs. Story 2	Rank order correlation	
	Men	Women
Usage as Old-Pairs (12 pair)	.929*	.943*
Usage as Old-Pairs (8 pair)	.818*	.883*
Usage as Single-Words (12 pair)	.800*	.961*
Usage as Single-Words (8 pair)	.441†	.925*
Usage as New-Pairs (12 pair)	.716*	.846*
Usage as New-Pairs (8 pair)	.553†	.858*

* Correlations are significant at 0.01 level.
† Correlations are significant at 0.05 level.

TABLE VI

DIFFERENT WORD-USAGE IN STORIES 1 AND 2

Word-List	Times member used					
	as Old-Pair		as Single-Word		as New-Pair	
	Total both stories	% Diff.	Total both stories	% Diff.	Total both stories	% Diff.
Men:						
12 pair	466	42.1	1320	41.4	463	85.7
8 pair	438	39.3	614	58.3	282	87.2
Women:						
12 pair	627	46.1	2049	40.1	696	87.1
8 pair	592	43.9	968	54.8	409	87.8

If individual words were used with similar frequency as Old-Pairs, Single-Words, and New-Pairs in the two stories, is this because *S*s write similar stories and thus merely repeat themselves? To answer this question, the number of times members of pairs were used as Old-Pairs, Single-Words, and New-Pairs in the two stories were combined for each individual and the number of duplicates and differences in usage noted.

The results for the men, based both on the 12-pair and 8-pair word lists, are shown in the upper half of Table VI. With respect to the 12-pair word list, we find that, on the average, men used different Old-Pairs 42.1% of the time; different Single-Words 41.4% of the time; and different New-Pairs 85.7% of the time. With respect to the 8-pair word list, the proportion of different words used for each classification is very much the same as for the 12-pair list. Although the 8-pair list might be expected to increase the variability of Single-Word usage, the increase is not significant.

The results for the women, reported in the lower half of Table VI, show almost identical patterns with respect to the degree of variation in word-selection for Stories 1 and 2, as well as with respect to the degree of similarity in results for the 12- and 8-pair analyses.

The small number of times that New-Pairs were repeated by both men and women supports the point that although certain pairs tended to be recombined, these recombinations usually were different from one story to the next. The data support the conclusion that there is a tendency in story writing to avoid using the words in the same way. Although standard learning concepts can account for repetition in behavior, the high degree of variability shown in the two stories (about 40% for Old-Pairs and over 85% for New-Pairs) is a phenomenon that indicates the need for additional concepts. Actually the degree of difference, especially for New-Pair usage, suggests that repetition is actively avoided.

Comparison of Stories 1 and 2. Since the Neutral pairs tended to be used less often as pairs, their inclusion masks individual differences in the usage of Old-Pairs and New-Pairs. We therefore have excluded them in obtaining individual scores on the usage of Old-Pairs, Single-Words, New-Pairs, and the total number of Experimental-Words. The upper half of Table VII presents the results for men and the lower half the results for women. This Table shows the means, standard deviations, and range of the various types of word usage for Stories 1 and 2. The similarity in scores for the two stories and for both sexes is very striking. The mean number of words used in Story 1 and Story 2 (Story-Length) is 185.1 and 178.8, respectively, for men, and 206.3 and 204.4, respectively, for women. Comparison of sexes shows women writing longer stories, the first being significant at the 0.10 level, the second at the 0.05 level. With respect to the Experimental-Word score, no story or sex differences were obtained. (Although there were only 16 words available, this maximum could be exceeded because both words were counted for each New-Pair created. For example, *damp weather* and *damp crackers* would be scored as an Old-Pair and a New-Pair, respectively, for a total of four Experimental-Words used without duplication. We also determined scores based upon duplicated usage, but found that this measure did not alter the relationships.) The similarities in the number of Experimental-Words used in the two stories by the two sexes is of special interest, because the number does not increase for the second story as one might expect from the effects of practice.

With respect to the use of Old-Pairs, there is a significant drop from Story 1 to Story 2 for both sexes. The means for Single-Words show, however, no significant change. The usage of New-Pairs shows a slight increase, which is not significant for the men but is significant at the 0.10 level for the women.

The last line of Table VII shows the Difference-Score (obtained by subtracting the New-Pair score from the Old-Pair score) which also shows a significant difference in the means of the two stories for both sexes. This measure of the two stories reflects the less frequent use of Old-Pairs and the modest increase in New-Pairs. Individual variability in this score is indicated by the large standard deviation.

TABLE VII

COMPARISON OF WORD-USAGE IN STORIES 1 AND 2
(Based on 8 Pairs)

Scores	Mean		*SD*		Range	
	Story 1	Story 2	Story 1	Story 2	Story 1	Story 2
Men:						
Story-Length	185.1‡	178.8§	78.76	65.66	43 to 439	51 to 370
Experimental-Words	15.9	14.9	5.18	5.24	3 to 35	0 to 32
Old-Pairs	3.7*	2.8*	2.10	2.15	0 to 8	0 to 8
Single-Words	4.4	4.8	2.82	3.19	0 to 11	0 to 11
New-Pairs	2.0	2.2	1.97	1.76	0 to 10	0 to 7
Difference-Score	1.7*	.6*	3.31	2.99	−7 to +8	−6 to +8
Women:						
Story-Length	206.3‡	204.4§	75.08	73.96	64 to 381	43 to 374
Experimental-Words	16.1	15.3	5.24	4.26	10 to 32	4 to 31
Old-Pairs	3.4*	2.7*	2.19	2.00	0 to 8	0 to 8
Single-Words	5.0	5.1	2.57	2.68	0 to 12	0 to 12
New-Pairs	1.9†	2.3†	1.89	2.04	0 to 8	0 to 10
Difference-Score	1.5*	.4*	3.63	3.49	−6 to +8	−9 to +8

* Difference between Stories 1 and 2 significant at 0.01 level.
† Difference between Stories 1 and 2 significant at 0.10 level.
‡ Difference between men and women significant at 0.10 level.
§ Difference between men and women significant at 0.05 level.

Individual differences. To reveal the extent of individual differences, the scores for the two stories were combined. Since the distributions of the men and women were so much alike, we included all the *S*s in plotting the frequency-distributions.

Fig. 1 shows the frequency distribution of the Old-Pair scores based upon 8 pairs of words (Neutral pairs excluded). Of interest is the fact that as many as 7 *S*s used no Old-Pairs at all, while 7 used 15 or more. The most frequent score was 8, which was exceeded by 41 *S*s but was not reached by 99.

The distribution for Single-Word scores is shown in Fig. 2. Because the Neutral pairs (fragmented because of situational factors) were excluded, this distribution represents primarily individual differences in preference for this usage. Even with these pairs eliminated, the scores range from 0 to 21, with a mean of 9.7.

Fig. 3 shows the frequency distribution for New-Pairs. Here we find 21 *Ss* making a zero-score, while 8 make a score of 11 or more. This distribution shows the tendency to use a large number of New-Pairs to be confined to a few *Ss* and is more skewed than the distribution for Old-Pairs. As few as 12.8% of the *Ss* created 32.6% of the New-Pairs, while 12.9% contributed none at all.

If the tendencies toward the use of Old-Pairs and the use of New-Pairs are in conflict, then the Difference-Score should yield a still broader distribution, whereas if these usages represent similar tendencies, the distribution should be more central. Subtracting the New-Pair scores from the Old-Pair scores should yield negative scores if a preference for New-Pairs dominates and a positive score if the dominant preference is for Old-Pairs. Fig. 4 shows the distribution for the Difference-Scores to range from —13 to +16. It is apparent that the range of scores has almost doubled and that the extreme scores have changed very little.

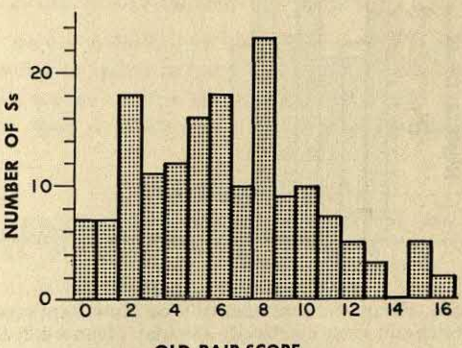


FIG. 1. FREQUENCY DISTRIBUTION OF OLD-PAIR SCORES
Two stories and sexes combined—8 pairs. Mean = 6.3; *SD* = 2.62

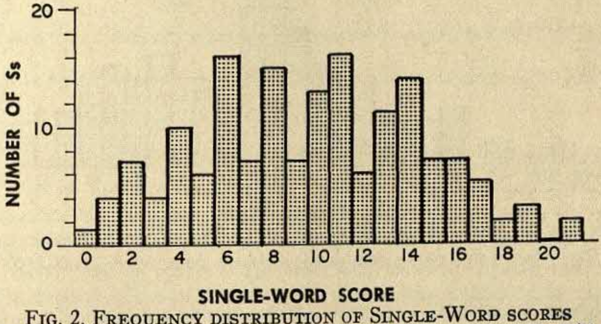


FIG. 2. FREQUENCY DISTRIBUTION OF SINGLE-WORD SCORES
Both stories and sexes combined—8 pairs. Mean = 9.7; *SD* = 4.71

Usage of Old-Pairs seems to be the predominant tendency in the group, which is revealed by the fact that 55.2% of the *Ss* made positive scores and 33.1%

made negative scores. Which tendency predominates, however, varies greatly from *S* to *S*.

Since the Difference-Score sets the usage of Old- and New-Pairs in opposition to each other, it serves as a good measure of any individual changes in word usage between Story 1 and Story 2. Table VIII shows an increase in the number of persons having negative Difference-Scores. The Chi-square test of the change in the frequency of positive to negative Difference-Scores is significant at the 0.01 level of confidence for the men and at the 0.05 level for the women.

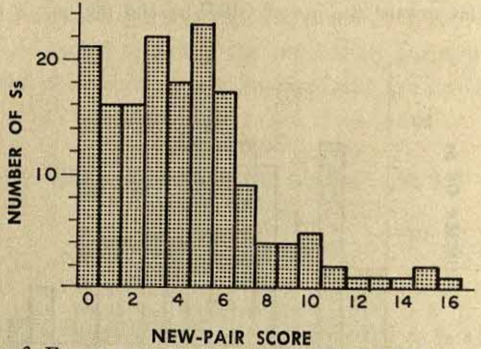


FIG. 3. FREQUENCY DISTRIBUTION OF NEW-PAIR SCORES
Two stories and sexes combined)—8 pairs. Mean = 4.2; *SD* = 3.34

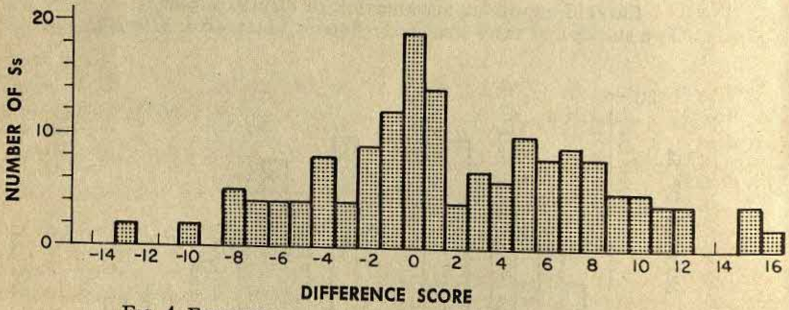


FIG. 4. FREQUENCY DISTRIBUTION OF DIFFERENCE-SCORES
Two stories and sexes combined—8 pairs. Mean = −2.1; *SD* = 6.02

The relative increase in negative Difference-Scores reveals that experience in writing stories around the experimental words causes an increase in the prevalence of New-Pairs. Thus, familiarity with the words learned tends to increase the formation of new combinations and decrease the usage of familiar pairs.

Relationships between measures. How do the various scores on the usages of words relate with one another? Table IX shows the correlation matrix for the word usage measures for Stories 1 and 2 as well as for men and women. These relationships are based upon the 8 pairs. Although the

Neutral pairs have been excluded, this fact has not affected the correlations, which we found to be of similar magnitude.

Comparison of men and women coefficients reveals striking similarities. Of the correlation coefficients in Story 1, 11 were significant for the men (0.05 level of confidence or better) and 11 for the women; 10 of these were the same measures. Of the 15 correlation coefficients in Story 2, 8 were significant for the men and 11 for the women, and these 11 included all of the measures that were significant for the men.

Comparison of Story 1 with Story 2 shows that only 3 significant relationships found in Story 1 failed to appear in Story 2. Furthermore, all 8 of the relationships that were significant for both men and women (a total of 16) in Story 2 also were significant for both sexes in Story 1. The consistent similarity of word usage in the two stories reveals that the way our *Ss* utilized learned information is a stable, individual trait in that the same behavior pattern in word-usage appears in both stories, regardless of content. As previously shown in Table VI, the differences that do appear involve a general tendency to reduce the number of Old-Pairs in the second story. The 8 relationships that hold for both stories and both sexes warrent consideration. One of these is the relationship between Story-Length and the Single-Word score. Persons who wrote the longer stories used more Single-Words, but they did not use more Old-Pairs or more New-Pairs (3 New-Pair relationships approach zero and 1 is positive).

TABLE VIII

PREFERENCES FOR OLD-PAIRS VS. NEW-PAIRS

Stories	Percentage of scores showing		
	positive Diff.	zero Diff.	negative Diff.
Men ($N=67$):*			
Story 1	65.7	13.4	20.9
Story 2	43.3	14.9	41.8
Women ($N=96$):†			
Story 1	62.5	6.2	31.3
Story 2	44.8	14.6	40.6

* Chi-square test of difference between stories for the men is 7.80, 2 *df*, significant at 0.01 level.

† Chi-square test of difference between stories for the women is 7.18, 2 *df*, significant at 0.05 level.

Three of the consistent relationships above-mentioned involve the Experimental-Word score, which is found to be related to all three types of word usage. The fact that the Difference-Score shows negative relationships with the Experimental-Word score (one of which is significant at the 0.05 level) suggests that the Difference-Score is a measure quite distinct from either the Old- or New-Pair scores from which it is derived. It appears that the positive relationships between Story-Length and the three types of word-usage, as well as those between Experimental-Words and word-usage, is caused by the mere fact that all three forms of word-usage contribute to Story-Length and the Experimental-Word score. The tendency to write long stories or the ability to use experimental words, therefore, cannot be regarded as a trait contributing to any particular kind of word-usage.

The relationship between Old-Pair, Single-Word, and New-Pair scores are of most interest since they measure the differences in word-usage. Here we find Old-Pairs to be significantly (but negatively) related to Single-Word scores for both stories and both sexes. Persons who are most prone to use Old-Pairs are least prone to fragment the pairs and hence use fewer Single-Words. This is the fifth of our stable relationships. With the stories combined, the correlation coefficients for Old-Pairs vs. Single-Words is —0.407 for men and —0.469 for women.

TABLE IX

Intercorrelations Between Measures of Word-Usage

Word-usage	Sex	Experimental-Words	Old-Pairs	Single-Words	New-Pairs	Diff.-Score
Story 1:						
Story-Length	M	.572*	.175	.373*	.323*	−.097
	W	.345*	.144	.415*	.007	.083
Experimental-Words	M		.375*	.393*	.630*	−.153
	W		.241†	.449*	.324*	−.023
Old-Pairs	M			−.390*	−.295†	.811*
	W			−.391*	−.578*	.905*
Single-Words	M				.239†	−.403*
	W				.298*	−.392*
New-Pairs	M					−.794*
	W					−.870*
Story 2:						
Story-Length	M	.166	−.072	.392*	−.021	−.039
	W	.231†	.048	.268*	−.033	.042
Experimental-Words	M		.459*	.354*	.608*	−.028
	W		.217†	.397*	.582*	−.220†
Old-Pairs	M			−.410*	−.166	.815*
	W			−.379*	−.472*	.853*
Single-Words	M				.121	−.365*
	W				.155	−.310*
New-Pairs	M					−.707*
	W					−.861*
Stories combined:						
Old-Pairs	M			−.407*	−.274†	
	W			−.469*	−.537*	
Single-Words	M				.292†	
	W				.389*	

* Significant at 0.01 level.
† Significant at 0.05 level.

With regard to Old-Pairs and New-Pairs, we also find four negative relationships, three of which are significant. When the stories are combined, the correlation coefficients for Old-Pairs vs. New-Pairs are —0.274 for men ($p < 0.05$) and —0.537 for women ($p < 0.01$). The tendency to use Old-Pairs, therefore, is also found to be in conflict with the tendency to make up New-Pairs.

The question of the relationship between Single-Words and New-Pairs now becomes critical, since each is negatively correlated with Old-Pairs. All four of the correlations are positive, but only one is significant and all these relationships are less than 0.3. When the stories, however, are combined, the relationship between Single-Words and New-Pairs becomes significant (see last line of Table IX). It appears, therefore, that the Single-Word score and the New-Pair score primarily are measures of different functions. The processes of fragmenting and regrouping are not in conflict and seem to overlap to some degree.

Other relationships. Three relationships remain that are significant and present in both sexes and both stories. All of them concern the relationships between the Difference-Score and the three types of word usage measures. Since the Difference-Score is derived by subtracting New-Pairs from Old-Pairs, the large positive correlation with Old-Pair scores and the large negative correlation with New-Pair scores are to be expected. Because Single-Word scores are slightly related to New-Pairs and negatively related to Old-Pairs, the relatively small but negative correlation is also to be expected.

Since scores for the usage of Single-Words and New-Pairs show a small but positive correlation, we attempted to explore the nature of this relationship. The scattergram shown in Fig. 5 reveals that *S*s scoring well above average on Single-Words tend to score either low or high on New-Pairs, whereas those who score well above average on New-Pairs make relatively high Single-Word scores.

To test this relationship further, we selected the high scoring individuals on each type of word usage and examined their scores on the other types of word usage. Table X shows the results of this analysis. The top third of the table shows that the 22 *S*s who make the highest Old-Pair scores make scores below the means on Single-Words and New-Pairs; the percentage of those making scores above the means of the total population being 18.2 and 0, respectively. Since Old-Pairs are negatively correlated with both Single-Words and New-Pairs, the percentages of those scoring above their means should be below 50%. The fact that this occurred indicates that the high-scoring individuals have contributed materially to the negative correlations.

The middle section of this table shows the results for the 25 *S*s making the highest Single-Word scores. The fact that less than half (20.0%) of these make scores above the Old-Pair mean reflects the negative correlation between Old-Pairs and Single-Words, but the fact that only 56.0% make scores above the mean for New-Pairs fails to support a positive correlation between Single-Words and New-Pairs. An examination of the individual scores for these 25 *S*s shows that 19 make

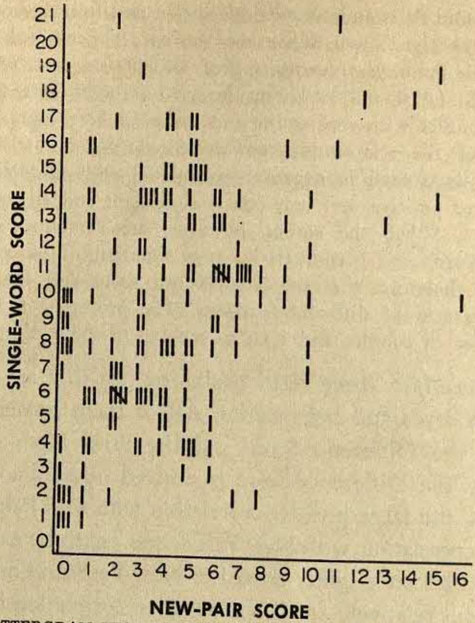


FIG. 5. SCATTERGRAM SHOWING RELATIONSHIP BETWEEN SINGLE-WORD AND NEW-PAIR SCORES
Two stories and sexes combined—8 pairs.
Mean single-words = 9.7; Mean New-Pairs = 4.2

TABLE X

NATURE OF RELATIONSHIPS BETWEEN WORD-USAGES

Word-usage	Mean, select group	Range	Over-all mean	% group select scoring above over-all mean
Ss, highest Old-Pair scores (N=22):				
Old-Pairs	12.9	11 to 16	6.3	100.0
Single-Words	5.1	0 to 17	9.7	18.2
New-Pairs	.9	0 to 4	4.2	0.0
Ss, highest Single-Word scores (N=25):				
Old-Pairs	4.5	1 to 11	6.3	20.0
Single-Words	16.8	15 to 21	9.7	100.0
New-Pairs	5.6	0 to 15	4.2	56.0
Ss, highest New-Pair Scores (N=21):				
Old-Pairs	4.2	0 to 9	6.3	28.6
Single-Words	13.0	7 to 21	9.7	90.5
New-Pairs	11.2	8 to 16	4.2	100.0

New-Pair scores from 0 through 5; 5 make scores from 11 through 15; and only 1 makes a score (9) that falls between these clusters. Thus *Ss* making high Single-Word scores reveal a bimodal distribution of New-Pair scores.

The lower third of Table X shows the results for the 21 *Ss* making high New-Pair scores. The fact that only 28.6% of them make above average Old-Pair scores reflects the negative correlation between New-Pairs and Old-Pairs, and the fact that 90.5% of these *Ss* score above the mean on Single-Words ($p < 0.01$) clearly reflects the existence of a positive correlation between New-Pairs and Single-Words. It appears, therefore, that *Ss* who do best with New-Pairs also are proficient in using Single-Words, but *Ss* who are proficient in using Single-Words are not necessarily proficient in using New-Pairs. This one-way relationship indicates that forming New-Pairs includes fragmenting, but that fragmenting does not indicate a tendency toward reorganization.

TABLE XI

ORIGINALITY RATINGS AND WORD-USAGE

Word-usage	Originality rating	
	Men	Women
Story 1:		
Experimental-Words	.139	.002
Old-Pairs	.169	−.193
Single-Words	.144	.241*
New-Pairs	−.078	.077
Story 2:		
Experimental-Words	−.167	.029
Old-Pairs	.005	−.026
Single-Words	.184	.099
New-Pairs	−.421†	−.070

* Significant at the 0.05 level.
† Significant at the 0.01 level.

Relation between word-usage and story-ratings. It will be recalled that each story was rated in terms of originality (using the Getzels-Jackson procedure) and that the two stories were compared and rated on a 5-point scale with respect to dissimilarity of (1) topic and (2) direction. Since the correlations of the judgments of the two raters were high, we combined their ratings in making this analysis of story originality and word-usage. These results are shown in Table XI. Only two significant relationships are found in this analysis, one for men in Story 2 and one for women in Story 1. Neither of these relationships repeats itself in the two stories or appears for both sexes.

Scores based upon 12 rather than 8 pairs were also computed. It was found that the significant correlation for the men remains, while the significant correlation for the women shifts from Single-Words to Old-Pairs (0.05 level of confidence).

None of the correlation coefficients obtained suggests that originality is related to word-usage in any consistent manner, and the obtained relationship for men in the second story is opposite in direction to what one might reasonably expect.

Table XII shows the relationships between word-usage in Stories 1 and 2 and each of the two measures of dissimilarity between stories. (It will be recalled that the topic of the first story was specified by the *E*s; that of the second story was determined by the *S*.) Only 3 of the 32 correlation coefficients are significant, but none of these holds for both men and women or for both stories. The negative relationships between the New-Pair score for men on Story 2 with both measures of dissimilarity of stories implies that New-Pair users write more similar stories, while the one relationship between Single-Words and one measure of dissimilarity, though positive, is significant at the 0.05 level only. Thus Table XII also fails to reveal any meaningful relationships between word-usage and both measures of dissimilarity.

TABLE XII

DISSIMILARITY OF STORIES AND WORD-USAGE

Word-usage	Dissimilarity Rating between Stories			
	topic		direction	
	Men	Women	Men	Women
Story 1:				
Experimental-Words	.048	.084	.102	.062
Old-Pairs	.088	−.057	.166	−.016
Single-Words	.061	.106	.009	.168
New-Pairs	−.111	−.024	−.083	−.106
Story 2:				
Experimental-Words	−.070	−.004	−.024	.036
Old-Pairs	.025	−.062	.091	−.031
Single-Words	.213	.131	.252*	.195
New-Pairs	−.328†	−.076	−.376†	−.098

* Significant at 0.05 level.
† Significant at 0.01 level.

The fact that the New-Pair scores for men is negatively correlated both with the dissimilarity of their stories and with their originality scores in Story 2 suggests that these ratings may be related. Table XIII shows that the relationship between the originality rating for Story 2 and the dissimilarity ratings are indeed positively correlated. Ten of the 12 correlation coefficients are significant at the 0.01 level of confidence.

In an attempt to explore further possibilities of relationship between word usage and story ratings, we compared stories of the 6 *S*s making the highest negative Difference-Scores with the 6 making the highest positive Difference-Scores. No differences in story originality and dissimilarity revealed themselves. We also selected the 10 stories receiving the highest originality ratings. The individuals writing these stories showed no unique patterns of word-usage scores.

Relation between word-usage and external measures of mental compe-

tence. In addition to the measures described thus far, we had available a number of test-scores from university records. These included: (1) Scholastic Aptitude Test scores (verbal); (2) Scholastic Aptitude Test scores (mathematical); (3) OAIS Achiever Personality scale; (4) OAIS Intellectual Quality scale, and (5) OAIS Creative Personality scale. In addition we administered the Guilford Gestalt Transformation Test and obtained learning scores based on the memory-checks.

It was thought that the verbal and mathematical measures of the Scholastic Aptitude Test (*SAT*) would be relevant because the story-writing activity required in our experiment was primarily verbal. The OAIS measures also are of interest because they include motivational and intellectual aspects of personality. The Guilford Gestalt Transformation Test was selected because it seemed to require a form of fragmenting activity and escape from functional fixity (*e.g.* "Which of the following is most useful to start a fire: fountain pen, onion, *pocket watch,* light bulb, bowling ball). Finally, the scores on the recall-tests were used as a measure of the degree of learning of the experimental list of paired words and to indicate whether differences in the familiarity with the words influenced usage.

TABLE XIII

INTERCORRELATION OF RATINGS OF ORIGINALITY AND DISSIMILARITY

Intercorrelation	Originality Rating Story 2		Dissimilarity Rating of Topic		Dissimilarity Rating of Direction	
	Men	Women	Men	Women	Men	Women
Originality Rating Story 1	.534*	.470*	.215	.414*	.200	.541*
Originality Rating Story 2			.656*	.607*	.683*	.657*
Dissimilarity Rating of Topic					.638*	.680*

* Correlations significant at 0.01 level.

Table XIV shows the correlations between these measures and the word-usage for both men and women. The first numbers in each column are the correlations with Story 1 and the numbers in parentheses are correlations with Story 2. The interesting feature of this Table is that all correlations are low and that none of them is significant for both men and women. Only 14 of the 168 correlation coefficients are significant, 9 for men and 5 for women. Only two instances occur in which the correlation coefficients are significant for both stories: the positive correlations of OAIS Intellectual Quality with New-Pairs and with Experimental-Words. Since our measures of word-usage are stable for both stories and sexes, the failure to obtain consistent correlations with external measures indicates that none of the external measures has a heavy loading of any of our highly specific measures of word-usage.

The measures of memory-scores and word-usage likewise showed relationships approaching zero. Only one out of 24 correlation coefficients was significant—that with Experimental-Words for men in Story 2. Thus it is obvious that the degree

TABLE XIV

CORRELATIONS BETWEEN EXTERNAL AND INTERNAL MEASURES

Measures	Sex	Story-Length	Experimental-words	Old-Pairs	Single-Words	New-Pairs	Diff.-Score
Scholastic Aptitude Test (Verbal)	M	.208(−.111)	.224(.120)	−.094(−.002)	.362†(.142)	.139(.048)	−.131(−.027)
	W	.097(.147)	.110(−.037)	.149(.291†)	−.158(−.007)	−.007(−.323†)	.091(.345†)
Scholastic Aptitude Test (Math)	M	.010(−.096)	−.132(−.091)	−.153(−.143)	−.196(−.095)	.191(.149)	−.197(−.189)
	W	−.004(−.145)	−.009(.025)	−.217(−.118)	.093(.107)	.221*(.080)	−.245*(−.112)
OAIS Achiever Personality	M	.012(.117)	−.006(−.179)	.027(−.135)	.116(.058)	−.112(−.155)	.071(−.005)
	W	.112(.143)	.181(.205)	.006(.168)	.112(.137)	.185(−.022)	−.096(.108)
OAIS Intellectual Quality	M	.259*(.180)	.343†(.459†)	−.027(.055)	.216(.319†)	.332†(.322†)	−.221(−.149)
	W	−.007(−.120)	−.036(−.142)	−.077(.042)	.064(−.124)	−.028(−.106)	−.031(.091)
OAIS Creative Personality	M	.061(.023)	.086(.138)	−.128(−.050)	.378†(.137)	.000(.140)	−.088(−.117)
	W	−.033(.020)	−.063(−.120)	−.039(−.076)	.008(.023)	−.032(−.056)	−.006(−.016)
Guilford Gestalt Transformation	M	.110(−.016)	.058(−.032)	.070(.067)	.221(−.137)	−.172(−.007)	.136(.052)
	W	.058(−.053)	.119(.044)	−.006(−.027)	.175(.081)	.028(.052)	−.018(−.046)
Memory Score	M	.177(.094)	.158(.332†)	.032(.112)	.113(.227)	.070(.152)	−.029(−.009)
	W	−.025(.033)	.130(−.030)	−.119(−.132)	.194(.169)	.111(−.003)	−.130(−.075)

* Significant at 0.05 level. † Significant at 0.01 level.
Note: First number in each column is correlation for Story 1; correlation for Story 2 is in parentheses.

of learning had little effect on the tendencies to use Old-Pairs, Single-Words, and New-Pairs.

To test further the possibility of relationships with the external measures, we classified our *S*s into those having positive, zero, and negative Difference-Scores on each of the stories. The mean scores of the external measures for each classification were then compared. No differences approaching significance were found.

Relationships between story ratings and external measures. Since our measures of word-usage failed to be related to the several external measures and the story-ratings, we tested the relationship between the ratings and the external measures. Table XV shows the several correlations obtained for both men and women. Only 5 out of 48 correlation coefficients were significant, 1 for the men and 4 for the women, but no relationship was significant for both men and women. The Guilford Gestalt Transformation Test showed the best relationships with the ratings, 3 out of 8 of which were significant.

TABLE XV

CORRELATIONS BETWEEN STORY-RATINGS AND EXTERNAL MEASURES

Measures	Sex	Originality		Dissimilarity	
		Story 1	Story 2	Topic	Direction
Scholastic Aptitude (Verbal)	M	.236	.166	.165	.082
	W	.035	.089	−.068	.113
Scholastic Aptitude (Math)	M	.172	.094	−.005	.058
	W	.058	−.071	.058	−.062
OAIS Achiever Personality	M	−.111	.003	.067	.051
	W	.052	.222*	.086	.056
OAIS Intellectual Quality	M	.214	−.042	.089	.023
	W	.206	.091	.095	.168
OAIS Creative Personality	M	.018	.027	.203	.083
	W	.335†	.058	−.050	.131
Guilford Gestalt Transformation Test	M	.426†	.150	.125	.068
	W	.090	.219*	.074	.214*

* Correlations significant at 0.05 level.
† Correlations significant at 0.01 level.

Relationship between external measures. The question of whether any of our external measures were intercorrelated naturally comes to mind. Table XVI shows the correlation coefficients for the six external measures on which we had data for both males and females. Of the 30 correlation coefficients obtained, only 4 were significant and none was significant for both sexes. Thus the external measures we used apparently measured different mental traits, yet none showed a meaningful relationship with our external measures.

DISCUSSION

The experimental findings have raised a number of issues of a controversial nature that require discussion. Our expectations and hypotheses have not been supported in certain respects, and it appears that the way organisms store and utilize information may be more complicated than usually is supposed. Our *Ss* were exposed to 12 pairs of words sufficiently often to insure that they had overlearning of the paired words; but because the order of the pairs was changed for each trial, the learning of associations between the pairs was limited. Thus, in essence, we caused small segments of information (pairs of words) to be stored in the organism in order to study how these strongly associated portions, compared with weakly associated portions, would be processed when a problem arose. A problem-situation was posed by requesting the *Ss* to write a creative story about a prisoner, using as many of the stimulus-words as possible. As in all problem-situations, the problem posed our *Ss* required them to select and utilize materials from their past experiences, but in this instance they were given a body of relevant experiences with which to work.

TABLE XVI

INTERCORRELATIONS OF EXTERNAL MEASURES

Measures	Sex	SAT Math	OAIS Achiever Personality	OAIS Intellectual Quality	OAIS Creative Personality	Guilford Gestalt Transformation
Scholastic Aptitude Test (Verbal)	M	.118	−.169	.353†	.107	.161
	W	−.014	.299†	.078	.199	.214
Scholastic Aptitude Test (Math)	M		−.333*	.132	−.024	.198
	W		.132	.181	−.156	.125
OAIS Achiever Personality	M			.025	.053	−.148
	W			.063	−.045	.130
OAIS Intellectual Quality	M				.330†	.009
	W				.090	.023
OAIS Creative Personality	M					.088
	W					.107

* Correlations significant at 0.05 level.
† Correlations significant at 0.01 level.

One of the experimental variables may be disposed of very easily. We had anticipated sex differences because problem situations usually reveal a variety of differences in performance.[10] Only one difference was found:

---

[10] Hoffman and Maier, *op. cit.*, 1966, 382-390.

Women wrote significantly longer stories. The striking similarity in the scores of word-usage, the range of individual differences, and the agreement in related and unrelated functions not only demonstrated the similarity of men's and women's responses to operations required in this study, but lent conviction to the reliability of the relationships between our varied measures.

How do various individuals make use of this learned material when confronted with a problem-situation for which the material may be relevant? Our results revealed that individuals differed greatly in the processes they used. (1) Some *S*s predominantly used the paired words as learned (Old-Pairs) and were therefore greatly influenced by the experimentally produced associative bonds. They often went out of their way to fit an Incongruous learned pair into the story. (2) Others predominantly used Single-Words. Apparently, for them, each of the members of the paired words was readily accessible without the other, so that the associative bond was easily fragmented. (3) Still others made more abundant use of New-Pairs by combining a member of one pair with that of another.

The first of these processes conforms with definitions of learning. The other two are related to reorganization and indeed are negatively correlated with the first process. The functions of fragmenting and recombining show low correlations, however, with each other and therefore cannot be regarded as measures of the same process. It appears that the tendency to reorganization includes fragmenting, but that the tendency to fragment does not predict the tendency to reorganize (see Fig. 5 and Table X). Thus fragmenting seems to be a process in itself whereas reorganization incorporates fragmenting as an essential aspect of reorganization. Whether or not reorganization is facilitated when information is previously fragmented remains to be investigated.

The fact that *S*s differed in the degree to which each of the above functions was exhibited raises the question of whether the process predominantly used represents differences in their abilities or differences in their preferences. In either case, the usage represents a stable trait in that every *S* used each function to almost the same degree in writing two stories and accomplished this without a great deal of duplication of the specific word-pairs. In general, there is a lesser tendency to use Old-Pairs in the second story than in the first, and a rather marked tendency to show variability in word selection rather than repetition in writing two stories. Thus the change in the proportion of the three types of word usages is in the direction of greater variability and New-Pair formation rather than in the repetition of New-Pairs previously used.

It would appear that all persons should be able to perform each of the three functions, since they are very simple operations. The differences, however, in the pattern of usage reflects personality differences. Variations in the degree of ability for each function would influence both preferences and usage. Further exploration of these functions is needed to differentiate the influence of ability differences and preferences.

The variability in individual behavior that was obtained is of special theoretical interest. Although the *Ss* followed a similar pattern of word-usage in the two stories, the elements of the pattern varied because the *Ss* tended to avoid a duplication of the specific word-pairs. Repetitive behavior is to be expected because of the reinforcement of successful responses, but why did the *Ss* sometimes avoid it and why did some of them prefer variability to a greater degree than others? It has been shown that brain injury decreases variability, and that emotional excitement and frustration increase repetitive behavior.[11] Harlow has dealt with the problem of variability by postulating a curiosity need.[12] Although variability is a phenomenon that has received attention, its conflict with repetitive behavior needs clarification. Since repetitive behavior can be understood in terms of reinforcement concepts, how can these same concepts account for variability? The mechanisms of fragmenting and reorganization may explain why an organism varies its responses from one trial to the next while the learning mechanism may account for its repetitive behavior. If an organism has different mechanisms at its disposal, qualitative differences in behavior cease to be contradictory.

We had anticipated a relationship between our reorganization and fragmenting measures (New-Pair and Single-Word scores) and story ratings (originality and dissimilarity between stories). Failure to obtain some degree of relationship raises questions about our measures of word-usage as well as about the method of rating the stories. It is possible that in telling our *Ss* to use as many experimental words as possible and to write an original story, we may have established conflicting goals. In any event, persons who wrote the most dissimilar or similar stories did not differ in word-usage. This is a fact that must be accepted even if it does not conform to known theory.

---

[11] G. V. Hamilton, A study of perseverance reactions in primates and rodents, *Behav. Monogr.*, 3, 1916, (No. 13), 1-65; J. A. Hamilton and I. Krechevsky, Studies in the effect of shock upon behavior plasticity in the rat, *J. comp. Psychol.*, 16, 1933, 237-253; I. Krechevsky, Brain mechanisms and variability: III. Limitations of the effect of cortical injury upon variability, *ibid.*, 23, 1937, 351-364; Maier, *Frustration: The Study of Behavior Without a Goal*, 1949, 25-76; The behavior mechanisms concerned with problem solving, *Psychol. Rev.*, 47, 1940, 43-58.

[12] H. F. Harlow, Motivational forces underlying learning, in *Learning Theory, Personality Theory, and Clinical Research: The Kentucky Symposium*, 1954, 36-53.

The lack of relationships between word-usage and various intellectual qualities likewise poses problems. If intellectual qualities, such as creativity and originality, do permit individuals to reorganize experience, why have our direct measures of reorganization failed to show relationships with some of them? It might be claimed that we did not use the 'proper' tests of intellectual and creative ability. While some of our measures of intellectual and creative ability are open to question, none of them even suggested meaningful relationships or offered encouragement.[13]

Another possibility is that creativity is made up of many specific abilities and each test of creativity, originality, or problem-solving measures a different specific ability or function. This possibility must be seriously considered and implies that creativity is not a general trait.

A third possibility is that creativity was not measured by any of the three types of word-usage investigated in this experiment. This possibility certainly must be entertained. If this conclusion is accepted, it follows that researchers should abandon the notion that the reorganization of experience is essential to creativity. We isolated and measured three different processes or ways in which organisms use learned information. One of these involves associative bonds, the other two involve the reorganization of experience. If we failed to measure creativity, then creativity must be a function of something other than the reorganization of experience. In this event the concept of creativity is in need of better definition.

Finally, it is possible that the usual tests and measures of creativity are measures of ability to perform difficult tasks. There is a common assumption that creativity is a complex process and therefore must be a measure of superior abilities. Scores on such tests, therefore, would be determined by task difficulty and a collection of difficult tasks would fail to be measures of a unique ability. If the abilities to fragment and to reorganize information are dominant in an individual, there is no reason to suppose that these activities are difficult for him. Fragmenting associations and bringing together weak or unassociated elements under the stress of a problem are processes which differ from the formation and retention of associative bonds.[14] Each could be distributed in the population independently. Tasks that can be solved by the use of any of the three processes would therefore fail to reveal the presence or relative strengths of the three basic processes.

---

[13] O. K. Buros, *Sixth Mental Measurements Yearbook,* 1965, 336-338, 706-708, 1470-1475.

[14] Maier, *op. cit.,* 1960, 135-172.

Most learning situations permit opportunities for generalizing, differentiating, and showing insight, so that an interaction between varied processes would not be revealed.[15]

Our failure to show a relationship between creativity and the reorganization of experience, therefore, need not exclude the possibility that fragmentation and the reorganization of experience are basic mechanisms in creativity. Since tests of creativity utilize difficult tasks, they may be weighted to measure difficulty rather than a unique process such as the reorganization of experience. Learning tasks can range from simple to difficult without threatening the concept of associative bonds. Similarly, reorganization problems can be made simple enough for rats and children to solve without violating the concept of fragmentation or a tendency to recombine elements to form new pairs.[16] The commonly-held notion that creativity is a higher mental process, therefore, may be a misleading one, since it causes one to emphasize the difficulty of the task rather than the differences in function that varied tasks may demand. People can differ in the degree to which they possess the qualitatively different mental functions and these differences in kind may influence the way they approach problem-situations. Thus one person may search his memory for a solution to a problem, while another makes up the solution by combining parts of several memories.

The point at issue is the need to isolate qualitatively different functions. Our investigation of word-usage shows that individuals store and process simple data differently. Whether the term creativity is relevant and should be used depends on whether we wish to use the term "creativity" to designate a complex activity or whether we wish to designate a function different from learning and association formation.

If fragmenting and the reorganization of associative elements are seen as qualitatively different from association, it becomes essential to explore conditions favorable to them. Familiarity with the material appears to promote fragmenting and reorganization. This is suggested by our evidence showing a decline in the use of Old-Pairs in the second story. Certainly there was no evidence that weak associative bonds favored fragmenting. If subsequent research supports this point, it follows that fragmenting associations and the formation of new relationships are promoted by overlearning. Since associative bonds are also strengthened by overlearning, it becomes apparent that fragmenting is not akin to forgetting. Thus it cannot be said that fragmenting and association-formation represent different degrees of plasticity.

---

[15] Wertheimer, *op. cit.*, 1959, 1-12.

[16] Maier, Reasoning in white rats, *Comp. Psychol. Monogr.*, 6, 1929 (No. 3), 1-93; Reasoning in children, *J. comp. Psychol.*, 21, 1936, 357-366.

Whether we have to recognize three types or ways of using stored material (respecting associative bonds, breaking associative bonds, and reorganizing elements to form new combinations) or whether these can be reduced to two requires more research. The evidence thus far indicates that qualitatively different functions exist and that individuals differ in the way in which they use learned data.

THE EFFECT OF SOPHISTICATION ON RATIO- AND DISCRIMINATIVE SCALES

By RICHARD K. EYMAN,
Pacific State Hospital, Pomona, California

Stevens notes that modern efforts to quantify subjective magnitudes have led to three types of scales—all three designed to show how sensation grows as a function of stimulus-values.[1] Category- and discriminative scales have generally been shown to be nonlinearly related to magnitude-scales. "Are we," it has been asked, "to be left with a purely operational view, accepting different psychophysical laws for different scaling-methods? Is there any way to bridge the gaps between the different kinds of scales? In the face of discrepancies like this, it is not so much a question of which is correct, but how to reconcile the differences. Until we do, we have not learned as much about the phenomena as we should know."[2]

A number of studies have verified that the magnitude-scale is non-linearly related to both categorical and discriminative scales.[3] All of these studies have shown that scales based on discriminative or categorical data are approximately logarithmically related to the stimulus-continuum, whereas ratio or magnitude estimates of stimuli generally yield a power function between the stimulus and psychological continua.

Some investigations did not find a nonlinear relationship between these two

---

* Received for publication September 24, 1966. Supported in part by Grant No. MH-08667 from the National Institute of Mental Health. This study was directed by Professor J. P. Guilford of the University of Southern California. My thanks for assistance are also due Dr. Norman Cliff. Computing assistance was obtained from Health Sciences Computing Facility, UCLA, sponsored by NIH Grant FR-3.

[1] S. S. Stevens, Ratio scales, partition scales, and confusion scales, in Harold Gulliksen and S. J. Messick (eds.), *Psychological Scaling: Theory and Application*, 1960, 49

[2] J. P. Guilford, Psychological measurement a hundred and twenty-five years later, *Psychometrika*, 26, 1961, 115.

[3] C. E. Helm, S. J. Messick, and L. R. Tucker, Psychological models for relating discrimination and magnitude estimation scales, *Psychol. Rev.*, 68, 1961, 167-177; Eugene Galanter and Samuel Messick, The relation between category and magnitude scales of loudness, *idem*, 363-372; H. Eisler, Empirical test of a model relating magnitude and category scales, *Scand. J. Psychol.*, 3, 1962, 88-96 (b); Magnitude scales, category scales, and Fechnerian integration, *Psychol. Rev.*, 70, 1963, 243-253 (a); B. Schneider and H. Lane, Ratio scales, category scales and variability in the production of loudness and softness, *J. acoust. Soc. Amer.*, 35, 1963, 1953-1961; S. S. Stevens and Miguelina Guirao, Subjective scaling of length and area and the matching of length to loudness and brightness, *J. exp. Psychol.*, 66, 1963, 177-186; W. S. Torgerson, Quantitative judgment scales, in Gulliksen and Messick (eds.), *Psychological Scaling*, 1960, 21-31.

classes of scales.[4] Rather a linear relationship was reported. In all but the Warren and Poulton study, however, the stimulus-range used was either too narrow or subjective and, consequently, was not reported as a true test of the relationship between these scales.

It is well known that in psychophysical judgments *S*s may give "biased" responses as a function of training, maturation, and other experimentally manipulatable factors.[5] Senders and Sowards found that the *S*s yield proportions of judgments in accordance with their expectations based on their knowledge of the experimental situation.[6] No study has systematically investigated the effect of the *S*'s knowledge of the experimental situation and general sophistication on the relation of either discriminative or magnitude values to the stimulus-continuum.

The purpose of this study is to compare the psychophysical functions of discrimination and magnitude scales for *S*s with a knowledge of these methods; *S*s of normal intelligence but who have no knowledge of scaling-methods; and *S*s institutionalized as mentally retarded. By investigating the two scales as a function of sophistication, it will be possible (1) to determine the effect of knowledge and ability on ratio-estimates of stimuli relative to the ability to discriminate differences, and (2) to evaluate the scales in terms of their over-all suitability for general use.

Method. Judgments of length of lines and heaviness of weights constituted the two tasks used in this study. Both modalities are considered "prothetic" by Stevens.

Judgment of length. Strips of white felt (1 cm. wide) were laid against a black felt wall. The strips were viewed two at a time and were next to each other in a horizontal direction. Four series of line lengths were judged by all the *S*s.

The four ranges of strip-lengths used are presented below as Series A, B, C, and D. Series A, B, and C were used in connection with pair-comparison judgments. Series D, representing the total range, was used with constant-sum judgments. Visual angles for some of the stimuli are given in parentheses.

In Series A, B, and C, the *S* judged which of the two strips of felt presented was longer. A pretest demonstrated that the length of lines had to be very close to one another for *S* to have a discriminative problem as discriminal dispersions were generally small. Thus, because of the large number of comparisons needed to include even a limited range of the stimuli, the following design was used which entailed judging only those pairs of stimuli close together in magnitude.

---

[4] M. Björkman, An experimental comparison between the method of ratio estimation and pair comparison, *Rep. Psychol. Lab.*, University of Stockholm, 71, 1959 (a); G. Ekman and T. Kunnapas, Scales of conservatism, *Percept. mot. Skills*, 16, 2, 1963, 329-334 (c); G. Goude, *On Fundamental Measurement in Psychology*, Almqvist and Wiksell, 1962, 110-113; A Parducci, Range-frequency compromise in judgment, *Psychol. Monogr.*, 77, No. 2, 1963, 44-45; R. M. Warren and E. C. Poulton, Ratio and partition judgments, this JOURNAL, 75, 1962, 109-114.

[5] G. Ekman and L. Sjörberg, in P. R. Farnsworth, O. McNemar, and Q. McNemar (eds.), *Annual Review of Psychology*, 16, 1965, 451-474; Parducci, *op. cit.*, 1963; S. S. Stevens, Problems and methods of psychophysics, *Psychol. Bull.*, 55, 1958, 177-196.

[6] J. F. Corso, A theoretico-historical review of the threshold concept, *Psychol. Bull.*, 60, 1963, 356-370.

Stimulus no.	Series A	Series B	Series C	Series D
1	10.0(1°30′)	30.0(4°30′)	90.0(13°28′)	4.0(36′)
2	10.1	30.2	90.4	10.0
3	10.2	30.4	90.8	10.4
4	10.3	30.6	91.2	10.8
5	10.4	30.8	91.6	30.0
6	10.5	31.0	92.0	30.8
7	10.6	31.2	92.4	31.6
8	10.7	31.4	92.8	90.0
9	10.8(1°38′)	31.6(4°44′)	93.2(13°56′)	91.6
10				93.2
11				140 (20°50′)

Each of the three portions of the total range (Series A, B, C) of length was presented according to the experimental method shown in Table I.

Each 'X' in Table I represents a comparison between two stimuli. Note that Xs below the diagonal involve the same comparison as their counterparts above the diagonal. A total of 15 comparisons was needed for each of Series A, B, and C, using this procedure. Forty-five comparisons (3 x 15) were needed to evaluate the scales for length over the total range. The stimuli were presented in a random order with the exception that no stimulus was presented on successive comparisons.

TABLE I

Comparisons Required for Stimulus Series A, B, and C

	Stimuli K								
Stimuli J	1	2	3	4	5	6	7	8	9
1	—	×	×						
2	×	—	×	×					
3	×	×	—	×	×				
4		×	×	—	×	×			
5			×	×	—	×	×		
6				×	×	—	×	×	
7					×	×	—	×	×
8						×	×	—	×
9							×	×	—

In Series D, the *S* used the constant-sum method, dividing 100 points between the two-line strips of felt presented in pairs in accordance with his best judgment of the absolute ratio of the greater to the lesser. In presenting the stimuli, the following design was used.

Series D was presented according to the experimental method shown in Table II.

Each 'X' in Table II represents a comparison-judgment between two stimuli. The Xs below the diagonal involve the identical comparisons as their counterparts above the diagonal. Twenty-eight judgments were required for Series D during this procedure. The design presented was selected for scaling in order to economize on the number of judgments necessary, *i.e.* 28 instead of the 55 which would be required if all possible comparisons were made.

Four separate ranges of the length of lines had to be used to compare the two methods of scaling. The constant-sum method is suitable for a large range of stimuli required to investigate the questions under study. The method of pair-comparisons is only applicable to relatively small ranges of stimuli due to small discriminal dispersions for this type of task. Thus, for pair-comparison judgments for both length and weight, three ranges of stimulus-magnitude were selected *within* the total range sampled by the constant-sum method. The three limited ranges of stimulus-magnitude selected for use with the pair-comparison method sampled the low, middle, and upper range of the stimuli used in connection with the constant-sum method. Consequently, a total of four ranges of stimuli was needed for each modality.

The testing room measured 14 ft. 3 in. in length, 11 ft. wide, and 9 ft. high. *S* was placed 12 ft. 6 in. from the wall where the stimuli were presented. The lengths for Series A, B, and C were determined by a pretest using 10 undergraduate students in psychology and sociology. The total range and spacing between Ranges A, B, and C used for Series D was determined by previous studies with the exception that the first and last stimuli of the pair-comparison ranges were included in Series D.[7]

TABLE II

COMPARISONS REQUIRED FOR STIMULUS SERIES D

	Stimuli K										
Stimuli J	1	2	3	4	5	6	7	8	9	10	11
1	—	×			×			×			×
2	×	—	×	×	×		×	×		×	×
3		×	—	×							
4		×	×	—	×		×	×		×	
5	×	×		×	—	×	×	×		×	×
6					×	—	×				
7		×		×	×	×	—	×		×	
8	×	×		×	×		×	—	×	×	×
9								×	—	×	
10		×		×	×		×	×	×	—	
11	×	×			×			×			—

Judgment of weight. Four series of weights were also judged by all the *Ss*. As with length, the weights were presented two at a time using an apparatus suggested by Baker and Dudek.[8]

Like the lengths, the four ranges of weights were labeled A, B, C, and D, where A, B, and C were used for the pair-comparison judgments and D was used with the constant-sum method. The total range of weight sampled was proportionately identical to the length of lines. The weights in grams by series were:

---

[7] J. P. Guilford and H. F. Dingman, A validation study of ratio-judgment methods, this JOURNAL, 67, 1954, 395-410.

[8] K. E. Baker and F. J. Dudek, Weight scales from ratio judgments and comparisons of existent weight scales, *J. exp. Psychol.*, 50, 1955, 293-308.

Stimulus no.	Series A	Series B	Series C	Series D
1	100	300	900	40
2	103	305	910	100
3	106	310	920	112
4	109	315	930	124
5	112	320	940	300
6	115	325	950	320
7	118	330	960	340
8	121	335	970	900
9	124	340	980	940
10				980
11				1400

As with judgments of length, in Series A, B, and C, the weights were presented two at a time and *S* judged which was heavier. The identical design was used for weight as shown in Table I as the discriminal dispersions for weight were also small. Thus, 15 comparisons were needed for each of Series A, B, and C. Series D for weights followed the same procedure described for length in Table II. Twenty-eight judgments were required using this economical approach.

The stimuli were presented in a random order with the exception that no stimulus was presented on successive comparisons.

The weights for Series A, B, and C were determined by a pretest using the same *S*s as were used for the length of lines. Again, the total range and spacing was determined by Guilford and Dingman except that the first and last stimuli of ranges A, B, and C, were included in Series D. This allowed for a more direct comparison between the two methods of scaling.

Subjects. Three groups of subjects were used.

Group I. Twenty retarded *S*s were selected who were resident at Pacific State Hospital. All patients had IQs between 45–70, a diagnosis of undifferentiated or familial, and were over 21 yr. of age. They also had less than a 5th grade education. There were 11 men and 9 women. Their mean age was 30 yr. with a standard deviation of 4 yr.

These *S*s were administered two scales of the WAIS (arithmetic and vocabulary) to determine their potential ability for reporting the judgments required.[9] A minimum total scale score of 3.5 was used for the selection of patients. This was a very liberal minimum. It was by no means a significant criterion in ultimately determining who was capable of making consistent judgments. These measures were useful, however, as a control for ability in the evaluation of psychophysical judgments. The patients had a mean arithmetic scaled score of 3.5 and a standard deviation of 1.16. Their mean vocabulary score was 4.95 with a standard deviation of 1.81. The mean total score was 8.70, with a standard deviation of 2.58. The general norms for the WAIS scales were based on a reference group of 500 cases between the ages 20–34. For each of the two scales the average or mean score was 10 with a standard deviation of 3. The average total score (arithmetic and vocabulary) was 20 and a standard deviation was not specified for the combination of these scales.

Group II. Twenty normal *S*s were recruited from the psychiatric technicians employed by Pacific State Hospital to care for the patients. These individuals were relatively homogeneous with respect to education and general health. All had a high school education, good health (no handicaps) and had no acquaintance with scaling methods. Each *S* was carefully interviewed so that no one who had any college experience participated in the study. There were three men and seventeen women. Their mean age was 33 yr. with a standard deviation of 6 yr.

---

[9] David Wechsler, *WAIS Manual for the Wechsler Adult Intelligence Scale,* 1955, 37-38, 42-43.

These *Ss* also were administered the vocabulary and arithmetic scales of the WAIS in order to control for these abilities. The mean score of the arithmetic scale was 10.70, with a standard deviation of 2.25; the mean score on the vocabulary scale was 12.20 with a standard deviation of 2.31; and the total score on the scale was 22.9 with a standard deviation of 4.17.

Group III. Fifteen sophisticated *Ss* volunteered for the study; eight had Ph.D.- and seven had M.A.-degrees. All degrees were in psychology, sociology, or mathematics. All of these *Ss* were familiar with scaling and had at least one course on the subject. This group had 11 men and 4 women. Their mean age was 32 yr., S.D. 5 yr.

These *Ss* were not given the WAIS due to their familiarity with the test-materials and the homogeneity of their education which was well above that considered adequate for making the judgments used.

Procedure. After introductory instructions, all the *Ss* were given five practice periods for length and weight preparatory to making the pair-comparison judgments. Ten practice trials were given for the constant-sum method. For this method, broad limits of the correct answer were set up to determine whether or not *S* was capable of making the judgments. For example, two stimuli with a ratio of magnitudes of 55:45 would be considered correct within 50:50 to 75:25 limits. If an individual could not succeed in getting 10 correct practice trials within these limits, he was eliminated from the experiment. Thirty retarded patients were tested before 20 were found capable of ratio-judgments. Normal and sophisticated *Ss* had no difficulty at all. No test-stimuli were included in the practice series.

In addition, all the *Ss* were given careful instructions on lifting weights and judging line-lengths. They were allowed to lift weights as many times as they wished, but had to do it in a prescribed order. One *E* was used for all the *Ss* throughout the study. The presentation of the stimuli was carefully counterbalanced for all groups. Three conditions were controlled in this manner: (1) order of task, *i.e.* length or weight; (2) order of series, *i.e.* pair-comparison judgments (A, B, C) or constant-sum (D); and (3) order of stimuli, *i.e.* larger or smaller.

Scaling methods. Pair-comparison method. The method of pair-comparison was used to scale each task in terms of discriminative judgments. Gulliksen's least squares solution for incomplete data was used to scale the data.[10] This solution assumes Thurstone's Case V but still represents the most satisfactory analytical approach for an incomplete data matrix of the type used in this study.[11]

Goodness of fit tests discussed by Guilford and Torgerson were applied to check on the internal consistency of the solution.[12] As Eisler has shown, however, these tests are insensitive to the form of a possible latent distribution and measure only to what extent the distributions can all be normalized with some arbitrary transformation of the scale-axis.[13]

---

[10] Harold Gulliksen, A least squares solution for paired comparisons with incomplete data, *Psychometrika,* 21, 1956, 125-134.

[11] L. L. Thurstone, Psychophysical analysis, this JOURNAL, 38, 1927, 368-389.

[12] Guilford, *Psychometric Methods,* (2nd ed.), 1954, 160-168; Torgerson, *Theory and Methods of Scaling,* 1958, 185-190.

The relationship between length and weight stimuli and the associated values of the pair-comparison scale were investigated by considering two factors:

(1) The type of function describing the relationship between each series of lengths and weights and their respective scale-values for each of the three groups was determined by attempting to fit a number of well known functions and accepting the one giving the smallest standard error of estimate. The following functions were fitted to each of the 18 resulting psychophysical curves (3 groups × 2 modalities × 3 series of stimuli):

$$\text{(a)}\ \psi = a\,S + K$$
$$\text{(b)}\ \psi = a\,S^2 + b\,S + K$$
$$\text{(c)}\ \psi = a\,e^{bs} + K$$
$$\text{(d)}\ \psi = a\,\mathrm{Log}\,S + K$$
$$\text{(e)}\ \psi = a\,(S + K)^n$$

where

ψ = pair-comparison scale-values; S = stimuli; a, b, K, and n = constants determined by best fitting equation.

All of the above functions were fitted analytically and included the technique of a nonlinear least squares estimation.[14] Standard errors of estimate were furnished for each of the best fitting equations representing each of the five functions used.

(2) Since the ranges of stimuli covered within each series by the method of pair-comparison were very restricted, the psychophysical function would need to be investigated by evaluating the slopes at each of the three points of stimulus-range studied in Series A, B, and C. The values of the slopes of the psychophysical function at the three points over a large range of stimuli would be indicative of the type of function that exists. If the function is curvilinear, then the first derivative at any stimulus-value in Series A, B, or C would give an evaluation of the slope.

Specifically, if the over-all psychophysical function for the discriminative scale is linear, then the slope at stimulus-Series A would equal the slope at Series B and Series C. Two other possibilities exist. If 'a' is the slope and A, B, and C represent the low, middle, and high sections of the total range, then

$${}^{n}A > {}^{a}B > {}^{a}C,$$

would describe a negatively accelerated curve indicative of a power function

$$\Psi = a\ (5 + K)^n$$

where the exponent n is less than one or a logarithmic function

$$\Psi = \mathrm{Log}\ S + K.$$

---

[13] Eisler, Indirect scaling, *Rep. Psychol. Lab.*, No. 158, 1963 (d).

[14] H. O. Hartley, The modified Gauso-Newton method for the fitting of non-linear regression functions by least squares, *Technometrics*, No. 2, Vol. 3, 1961, 280-294.

However, if

$${}^{a}A < {}^{a}B < {}^{a}C,$$

then over-all function would have to be positively accelerated and therefore would likely be a power function as above, where the exponent n is greater than one or an exponential function

$$\Psi = ac^{bs} + K$$

The linear function already described would be

$${}^{a}A = {}^{a}B = {}^{a}C$$

which would satisfy the power function where the exponent 'n' is equal to exactly one. The quadratic function, due to its flexibility, could satisfy any of the previously mentioned possibilities. The quadratic function has, however, little theoretical value.

In this manner, it is also possible to evaluate the similarity of the discriminative scale based on three samples of the total range and the ratio scale covering the total range.

Meaningful comparison among the three ranges of stimuli for each task was achieved by use of 'standardized' measures which are results of a linear transformation involving translation (expected value) and rotation (variance), following Gulliksen's approach.[15]

Constant-sum method. The constant-sum method, as presented by Torgerson was used to scale each task in terms of ratio-judgments.[16] One exception should be noted. The geometric mean of the point assignments was used instead of the arithmetic mean customarily used. This procedure has been suggested by both Stevens and Coombs.[17]

Because of the problem considered by Pradham and Hoffman and Schiffman and Messick concerning averaging individual responses, an additional measure was taken to ensure the "representativeness" of the geometric mean of the individual point assignments.[18] It was desirable in this instance to exclude "blunders" or "gross errors" of an *S* on a specified trial from the geometric average of the point assignment for that trial. In this manner, the "efficiency" of this average and the degree to which it was representative of the individual point assignments would be improved. Dixon and Massey suggest a criterion for eliminating these extreme blunders on any given trial from the sample.[19] For samples of 14 to 30 observations the following formula is given:

$$r_{22} = \frac{X_3 - X_1}{X_{k-2} - X_1}$$

---

[15] Gulliksen, *op. cit.,* 125-134.

[16] Torgerson, *op. cit.,* 1958, 105-112.

[17] Stevens, The surprising simplicity of sensory metrics, *Amer. Psychol.,* 17, 1962, 29-39; C. H. Coombs, *A Theory of Data,* 1964, 371.

[18] P. L. Pradham and P. J. Hoffman, Effect of spacing and range of stimuli on magnitude estimation judgments, *J. exp. Psychol.,* 66, 1963, 533-541; H. Schiffman and S. Messick, Scaling and measurement theory, *Rev. Educ. Res.,* 33, 1963, 533-542.

[19] W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis,* (2nd ed.), 1957, 275-278.

where r_{22} is a statistic referred to a probability table for determining the level of significance regarding the chances out of 100 that the extreme value belongs to the same population of values.[20] (The criterion used to exclude values in this study was Type I error, $\alpha \leq 0.01$); x_i = a specified observation or judgment where the subscripts on the observations indicate that k observations have been arranged in order of size from the smallest to the largest and numbered

$$1 \text{ to } k,\ x_1 < x_2 \ldots < x_{k-1} < x_k.$$

Each value can then be tested.

In actual practice, no more than three judgments were excluded from the average point assignment on any trial. Approximately 35% of the trials had one to three extreme judgments removed on all *S*s tested.

Scaling was done on five subsets of the 11 stimuli used for each task. These subsets of stimuli were selected so that all 11 stimuli could be scaled even though an incomplete pair-comparison design was used (see Table II) in the interest of economizing on the number of judgments required. Using weight as an example for both tasks, the following sets of stimuli were scaled: (1) weights 40 gm., 100 gm., 300 gm., 900 gm., and 1400 gm.; (2) 100 gm., 124 gm., 300 gm., 340 gm., 900 gm., and 980 gm.; (3) 100 gm., 112 gm., and 124 gm.; (4) 300 gm., 320 gm., and 340 gm.; (5) 900 gm., 940 gm., and 980 gm. Subsets (3), (4), and (5) were poor choices, being too few stimuli to scale and resulting in extremely erratic scale-values in comparison with Subsets (1) and (2). Consequently, Subsets (3), (4), and (5) were discarded. Subset (1) included two anchor stimuli at the extremes of the stimulus-range employed and three weights in the middle range. The stimuli of Subset (2) represent the middle range and were selected to be identical to the two end stimuli used in Series A, B, and C for the pair-comparison approach.

The relationship between length and weight stimuli and the associated constant-sum scale-values was investigated in the same manner as previously discussed for pair-comparison scales.

The linear, quadratic, exponential, logarithmic, and power functions were fitted to combined scale-values. Although combining scale-values calculated separately from two submatrices of point-assignments can be a questionable approach, there was sufficient overlap in the stimuli scaled to produce very consistent scale-values. This was checked empirically before the scale-values determined from the subsets of stimuli were combined.

As with the pair-comparison scales, the criterion used for selecting the appropriate function was the standard error of estimate. Similarly, the power and logarithmic functions represented the preferred equations other factors being equal for theoretical reasons.

Investigating the slope of the best fitting function at specified points of the stimuli (ranges covered in Series A, B, and C using the pair-comparison method) offered a means of comparing the Thurstonian discriminative scale with the constant-sum scale. If the psychophysical function is linear, the slope is directly apparent. The slope of a curvilinear function can be evaluated simply by differentiating the scale (ψ) with respect to a specified stimulus-value (S). Appropriate comparison of the relationship between the pair-comparison and constant-sum scale values was investigated by this procedure. The internal consistency of the constant-sum method was checked by plotting observed derived ratios as a function of expected derived ratios determined from the scale-values. This was done for the three groups.

---

[20] Dixon and Massey, *op. cit.*, 1957, 275-278.

Results. The first question investigated was the type of psychophysical function characterizing the pair-comparison and constant-sum scales over the three levels of sophistication for length and weight.

TABLE III

PAIR-COMPARISON SCALE-VALUES AND ASSOCIATED CHI-SQUARE VALUES (GOODNESS OF FIT) FOR LENGTH OF LINES FOR PATIENT, NORMAL, AND SOPHISTICATE GROUPS

	cm.	patients	normals	sophisticates
Series A				
1	10.0	.000	.000	.000
2	10.1	.140	.143	.524
3	10.2	.245	.236	.670
4	10.3	.447	.592	1.011
5	10.4	.514	1.016	1.223
6	10.5	.616	1.428	1.796
7	10.6	1.298	2.277	2.222
8	10.7	1.190	2.555	2.755
9	10.8	1.644	2.805	3.036
		$X^2=3.397$ df=13	$X^2=0.631$ df=13	$X^2=5.599$ df=13
Series B				
1	30.0	.000	.000	.000
2	30.2	.122	.166	.180
3	30.4	.137	.616	.696
4	30.6	.369	.844	1.346
5	30.8	.583	1.018	1.629
6	31.0	.872	1.467	2.003
7	31.2	1.085	1.821	2.467
8	31.4	1.181	2.183	2.180
9	31.6	1.199	2.327	2.155
		$X^2=4.570$ df=13	$X^2=2.450$ df=13	$X^2=14.185$ df=13
Series C				
1	90.0	.000	.000	.000
2	90.4	.541	.163	.168
3	90.8	.554	.487	.516
4	91.2	.614	.785	.790
5	91.6	.296	1.318	.750
6	92.0	.293	1.462	1.185
7	92.4	.235	1.297	1.050
8	92.8	.280	1.397	1.104
9	93.2	.017	2.105	1.700
		$X^2=8.807$ df=13	$X^2=0.209$ df=13	$X^2=1.020$ df=13

Psychophysical function of the pair-comparison scale. Table III and IV present the pair-comparison scale-values and associated Chi-square values showing goodness of fit for length and weight of the three groups studied. Two findings are very apparent. The scale-values are somewhat unstable in a few instances, particularly in Series C for lengths. In other words, the scale-values do not always increase in value with an increase in stimulus-magnitude. Secondly, the X^2 values are all insignificant indicating that by the traditional criteria, Thurstone's Case V was appropriate and the "goodness of fit" of the model was adequate. Similar findings were apparent in Goude's attempt to use the pair-comparison method with lifted-weights.[21]

TABLE IV

PAIR-COMPARISON SCALE-VALUES AND ASSOCIATED CHI-SQUARE VALUES (GOODNESS OF FIT) FOR WEIGHT FOR PATIENT, NORMAL, AND SOPHISTICATE GROUPS

	gm.	patients	normals	sophisticates
Series A				
1	100	.000	.000	.000
2	103	.643	.570	.214
3	106	.706	.720	.301
4	109	1.136	1.014	.667
5	112	1.426	1.373	.493
6	115	1.439	1.754	.755
7	118	1.623	2.292	1.462
8	121	1.836	2.337	1.204
9	124	1.922	2.568	1.860
		$X^2=3.409$	$X^2=4.810$	$X^2=3.273$
		df=13	df=13	df=13
Series B				
1	300	.000	.000	.000
2	305	.249	.485	.548
3	310	.276	.443	.725
4	315	.200	.757	1.349
5	320	.400	.930	1.605
6	325	.534	1.172	1.570
7	330	.967	1.471	1.503
8	335	1.209	1.531	1.678
9	340	1.473	1.438	2.021
		$X^2=9.054$	$X^2=5.671$	$X^2=5.450$
		df=13	df=13	df=13
Series C				
1	900	.000	.000	.000
2	910	.245	.391	.106
3	920	.555	.380	.231
4	930	.180	.787	.374
5	940	.864	1.058	.781
6	950	1.000	1.191	.917
7	960	.944	1.285	1.232
8	970	.716	1.382	1.599
9	980	1.444	1.671	2.051
		$X^2=3.103$	$X^2=2.955$	$X^2=4.450$
		df=13	df=13	df=13

---

[21] Goude, *op. cit.*, 1962.

The matrices of joint proportions indicating the percentage of times one stimulus was chosen greater than another stimulus for Series A, B, and C for length and weight substantiated the scale-values in revealing that discrimination at times appeared to decrease as the stimuli got further apart. This causes the scale-values to decrease as a function of stimulus-magnitude and results from poor discrimination on the part of the *Ss*. Some of this difficulty could have been avoided if the stimuli were spread out more although the results of the pretest did not indicate such a possibility.

Table V presents the regression coefficients and standard error of estimates for the 18 best fitting linear equations. An attempt to fit the linear, quadratic, exponential, logarithmic, and power functions to these data revealed that in all instances, the linear equation provided the best fit or was at least as good as alternative best fitting functions. Since the scale-values were erratic in several instances and difficult to fit using any function, the linear equation was the preferred function—other factors being equal. The standard error of estimate was used as the criterion for evaluating the best fitting function.

TABLE V

REGRESSION COEFFICIENTS AND STANDARD ERROR OF ESTIMATES FOR EIGHTEEN LINEAR EQUATIONS FITTED TO PSYCHOPHYSICAL FUNCTIONS BASED ON PAIR-COMPARISON SCALES

Series	Length			Weight		
	Patients	Normals	Sophisticates	Patients	Normals	Sophisticates
Regression coefficients						
A	2.000	3.896	3.788	.074	.108	.071
B	1.858	1.521	1.556	.035	.038	.044
C	−.077	.609	.468	.015	.020	.025
Standard Errors of Estimates						
A	.158	.224	.130	.164	.120	.205
B	.091	.081	.312	.168	.135	.244
C	.212	.202	.155	.229	.105	.144

Table V shows clearly that a negatively accelerated curve describes the psychophysical function over the full range of stimuli studied. This finding holds for both length and weight for the retarded, normal, and sophisticated *Ss*. In other words, were it possible to tie together the scale-values for the three ranges studied (Series A, B, and C), it would be evident that as stimulus-magnitude increased, the slope of the curve would decrease. The flatter slopes shown for Series B and C in comparison to Series A illustrate this point. In general, the steeper the slope, the better the discrimination shown by the *Ss* involved.

It can be noted that generally the slopes of the lines are steeper for the sophisticate and normal groups than for the retarded group. This is particularly apparent in Table V concerning length. Consequently, the discrimination shown by the sophisticate and normal groups was usually superior to that shown by the retarded.

It can also be seen in Table V that the standard error of estimates are reasonably small. An analysis of variance suggested by Dixon and Massey was used to test for linearity of regression.[22] Only one function did not satisfy the requirements of linearity. This function was unique in having an essentially zero regression coefficient (length, patients, Series C).

TABLE VI

CONSTANT SUM SCALE VALUES FOR LENGTH OF LINES AND WEIGHTS FOR PATIENT, NORMAL, AND SOPHISTICATE GROUPS

	(cm. or gm.)	Patients (ψ)	Normals (ψ)	Sophisticates (ψ)
length	4	.127	.108	.162
	10	.274	.367	.395
	10	.185	.308	.348
	10.8	.281	.318	.411
	30	.982	1.200	1.103
	30	.866	.975	.981
	31.6	1.200	1.015	1.054
	90	3.131	3.480	2.964
	90	3.626	2.959	2.447
	93.2	5.112	3.485	2.761
	140	9.315	6.072	4.772
weight	40	.135	.123	.154
	100	.264	.285	.309
	100	.285	.270	.276
	124	.234	.289	.324
	300	.778	.844	.793
	300	.730	.856	.806
	340	.845	.956	.993
	900	3.639	3.063	4.117
	900	3.806	3.334	3.536
	980	6.383	4.702	3.959
	1400	9.906	8.561	6.446

Psychophysical function of the constant-sum scale. Table VI presents constant-sum scale-values for length and weight for each group. Scale-values are presented as computed from the two submatrices of point assignments, hence two scale-values are given for the same stimulus in three instances. As can be noted in Table VI, generally all scale-values are consistent even though calculated separately from two subsets of overlapping stimuli. The rationale for combining the scale-values in this manner was based upon the added stability of a psychophysical function provided by including scale-values associated with the two extreme stimuli (4- and 140-cm. lines; 40- and 1400-gm. weights).

---

[22] Dixon and Massey, *op. cit.*, 1957.

Good internal consistency was found for the constant-sum scales using plots of empirically derived ratios as a function of expected (derived) ratios determined from scale-values. Variability of the point assignments was by far the largest for the retarded *S*s (7-16 points), and smallest for the sophisticated, (1-6 points).

Five functions (linear, quadratic, exponential, logarithmic, and power) were attempted in fitting the constant-sum scales for length and weight to stimulus-magnitude. The power function in all instances was the best fitting function, although on two occasions (retarded *S*s) there was not much to choose between this function and an exponential or quadratic function. Still, as previously stated, other factors being equal, the power-function was the preferred function due to its widespread use and theoretical value. The standard error of estimate was used to evaluate the goodness of fit for all functions.

TABLE VII

EXPONENTS (n), SCALE CONSTANTS (a), ADDITIVE CONSTANTS (K), AND STANDARD ERROR OF ESTIMATES (SE) OF BEST FITTING POWER FUNCTIONS FOR LENGTH AND WEIGHT

$$\psi = a(S+K)^n$$

Subjects	Length				Weight			
	n	a	K	SE	n	a	K	SE
Patients	2.74	.0000050	51.40	.469	1.91	.0000085	73.04	.505
Normals	1.52	.0028214	15.03	.160	1.69	.0000394	6.71	.266
Sophisticates	1.37	.0044656	17.37	.158	1.30	.0005097	11.28	.160

Table VII presents the exponent (n), scale-constant (a), additive constant (K) and the standard error of estimates (SE) of the best fitting power-functions for these data. For both length and weight a very evident decrease in the value of the exponent is apparent, going from the retarded *S*s to the sophisticated. The standard error of estimates show satisfactory fits for all groups, although the power functions for the retarded were far less satisfactory than for the normal and sophisticate groups. It is also apparent from Table VII that all the exponents (n) are greater than one and hence describe a positively accelerated curve. The greater the exponent, the more positive acceleration present.

Comparison of the psychophysical functions for the pair-comparison and constant-sum scales. The relationship between the pair-comparison and constant-sum scales was investigated by comparing their respective psychophysical functions. This was done as previously discussed by comparing the slopes of the psychophysical functions at particular values of the stimuli, *i.e.* length and weight. Specifically, the regression-coefficients presented in Table V for the pair comparison-scales for three parts of the stimulus-range were compared with slopes of the associated power-functions within these same stimulus-ranges.

Slopes for the power-functions were calculated by differentiating the constant-sum scale ψ with respect to a specific value of the stimulus, S (length or weight). Thus:

TABLE VIII

SLOPES (FIRST DERIVATIVES) OF THE POWER FUNCTIONS IN TABLE VII FOR LENGTH AND WEIGHT AT SPECIFIED VALUES OF THESE STIMULI

Stimulus	Patients		Normals		Sophisticates	
	dψ/dS(CS)*	Regr. Coeff. (PC)†	dψ/dS(CS)*	Regr. Coeff. (PC)†	dψ/dS(CS)*	Regr. Coeff. (PC)†
Length:						
10 cm	.017	2.000	.023	3.896	.021	3.788
10.8 cm	.018		.023		.022	
30 cm	.028	.858	.031	1.521	.026	1.556
31.6 cm	.029		.032		.026	
90 cm	.074	−.077	.048	.609	.035	.468
93.2 cm	.077		.049		.036	
Weight:						
100 gm	.0018	.074	.0017	.108	.0028	.071
124 gm	.0020		.0019		.0029	
300 gm	.0036	.035	.0035	.038	.0038	.044
340 gm	.0040		.0038		.0039	
900 gm	.0087	.015	.0073	.020	.0053	.025
980 gm	.0094		.0077		.0054	

* CS = Constant-Sum Scale † PC = Pair-Comparison Scale

$$\psi = a\,(S + K)^{n}$$
$$d\,\psi/dS = na\,(S + K)^{n-1}$$

Table VIII gives the slopes of the power functions (constant-sum scales) listed in Table VII for length and weight within stimulus-ranges, Series A, B, and C. These slopes are listed adjacent to the slopes determined from the pair-comparison scales for these ranges of stimuli. The values of two slopes can be seen at the two ends of each range of stimuli listed in connection with the constant-sum scale. Since the power functions for the constant-sum scales change with increments in stimulus-magnitude, two estimates of the slopes were presented within each of the ranges listed.

It should also be pointed out that a direct comparison of the slopes connected with the constant-sum and the pair-comparison scale is impossible due to a different scaling factor for each of the scales. In other words, the absolute values of the slopes presented are contingent on the range of stimuli and associated scale-values. The range of scale-values for the pair-comparison scale is more "stretched out" than is true for the constant-sum scale. Comparisons between the slopes from the two scaling techniques must be based upon their relative change in value over the three ranges of length and weight investigated.

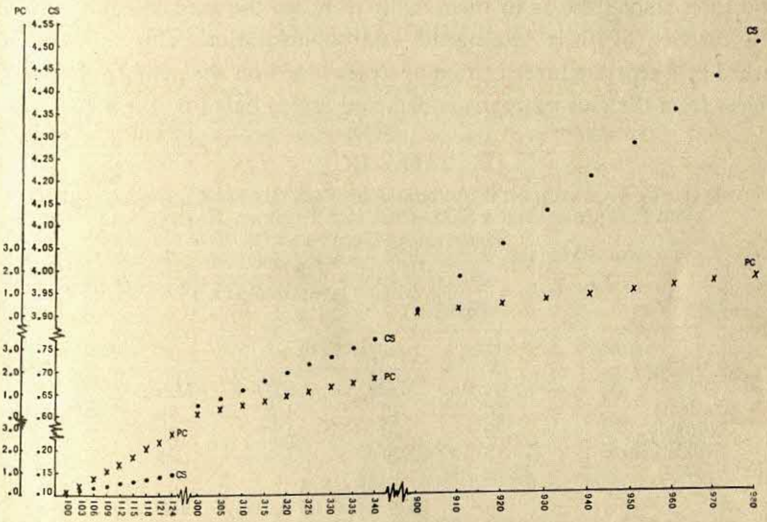


FIG. 1. COMPARISON OF PSYCHOPHYSICAL FUNCTION OF CONSTANT-SUM (CS) AND PAIR-COMPARISON (PC) SCALES FOR WEIGHT NORMALS

For all groups in both length and weight it can be seen from Table VIII that the values of the slopes increase for the constant-sum scale with increases in stimulus-magnitude. In contrast to this finding, the values of the slopes decrease for the pair-comparison scale with increases in stimulus-magnitude. The fact that the slopes of the power-functions increase with increases in stimulus-magnitude is obvious from the exponents of these functions described in Table VII. It is further apparent that the psychophysical functions for the pair-comparison scale is clearly different from the psychophysical functions connected with the constant-sum scale. The former is negatively accelerated and the latter is positively accelerated.

This finding is illustrated in Fig. 1 showing a comparison of the psychophysical function of the constant-sum and pair-comparison scales of the normal group for weight. The scale-values plotted in this figure consist of the "predicted" values based on the best fitting linear and power functions. Thus, the two scaling methods could be directly compared over Ranges A, B, and C for length and weight. Finally, a crude attempt was made visually to adjust for the different scaling factors discussed before.

Comparison of the accuracy of judgments of the retarded, normal, and sophisticated groups. A second goal of the study was to investigate possible differences among the *S*s in their ability to judge the magnitude of stimuli as a function of their intelligence and sophistication. This was accomplished by a separate investigation of scores based on the number of correct choices from the pair-comparison data and scores based on the accuracy of point assignments given in connection with the constant-sum scale. Table IX presents the results from an analysis of variance of the total number of correct choices given by an *S* with respect to the pair-comparison situation. Thus, each individual was assigned a score according to the number of correct choices he made from the 45 comparisons used in Series A, B, and C for the pair-comparison scale.

TABLE IX

ANALYSIS OF VARIANCE OF INDIVIDUAL ACCURACY SCORES CONNECTED WITH THE PAIR-COMPARISON SCALES FOR THE PATIENT, NORMAL, AND SOPHISTICATE GROUPS

* *S*s Score $= \sum_{j=1}^{45} C_{jk}$, where $C_{jk} =$ stimulus $k > j$

Subjects	N	Length		Weight	
		Mean	SD	Mean	SD
Patients	20	26	4.8	26	3.8
Normals	20	30	4.1	28	3.9
Sophisticates	15	29	3.6	28	2.8
		F=5.8 (Sig. at 0.05 level)		F=1.63 (Non-significant)	

* (a high score signifies good discrimination.)

The results of the analysis of variance of the pair-comparison data show a significant difference at the 0.05 level between group means for length. No significance was found between groups concerning weight. A glance at the group-mean scores shows the retarded *S*s have the lowest average of correct judgments for both tasks. Although the over-all significance of these results is questionable, they do support the psychophysical functions of the pair-comparison scales.

The accuracy of the point assignments from the constant-sum judgments were studied by constructing a difference-score for each individual. This score was simply the total difference in points over the 28 comparisons administered between the correct number of points describing the ratio between any two stimuli and the number that *S* gave. The higher the individual's score the poorer were his judgments of the 28 pairs of stimuli presented.

Table X gives an analysis of variance of the mean difference scores for the three groups. Significance was discovered between the average accuracy of judgments between groups for both length and weight. From the mean scores for each group it can be easily assessed that for both tasks the patients were by far the least accurate in their estimation of the stimulus-ratios and the sophisticates were by far the most accurate in their estimations.

TABLE X

Analysis of Variance of Individual Accuracy Scores Connected with the Constant Sum scales for the Patient, Normal, and Sophisticate Groups

*S*s Score $= \sum_{j=1}^{28} (V_{jk} - A_{jk})$, where V_{jk} = number of points out of 100 assigned to stimulus k when compared to stimulus j.
A_{jk} = actual number of points out of 100 specifying the ratio of stimulus k to stimulus j (a low score signifies accurate point assignments.)

	N	Length		Weight	
		Mean	SD	Mean	SD
Patients	20	379	75	384	94
Normals	20	157	126	218	106
Sophisticates	15	83	30	159	50
		$F = 53.1$ (Sig. at 0.001 level)		$F = 31.0$ (Sig. at 0.001 level)	

Summary of findings. The findings clearly show that the psychophysical functions for the constant-sum scales are different from those for the pair-comparison scales. The constant-sum scales were related to length and weight by curves concave upward (positively accelerated). The pair-comparison scales were related to length and weight by curves concave downward (negatively accelerated). The degree of upward curvature in the psychophysical functions for the constant-sum scales was inversely related to the *S*'s sophistication and was significantly related to errors in point-assignment. The derived negatively accelerated curves describing the relation between the pair-comparison scales and the stimulus-continua were also associated with *S*'s sophistication. In this instance, the steepness of the slopes describing the extent of the downward concavity of these psychophysical functions was generally less for the patients. The extent of this difference between groups for the pair-comparison data was far less than was found for constant data.

Discussion. The results of this study are relevant to arguments presented by Eisler and Torgerson.[23] Torgerson argues the two scales reflect two ways we have of judging relative differences between stimuli. Ratio-scales are most useful concerning quantitative relations among objects, relative gains, etc. Numbers are used to specify the ratio of one stimulus to another in such a way that errors are proportional to the differences between the stimuli judged. These errors increase exponentially as our "round number tendency" is such that the numbers are less precise and get rounder as we go up the scale. From results given here regarding length and weight, it was noted that overestimating increases with larger stimuli, and does so as a function of the *S*s' ability to use numbers—quite in line with Torgerson's argument.

The pair-comparison and category-scales are believed to be pure discriminative scales by both Eisler and Torgerson.[24] Eisler, following Stevens, states that the category-scale's deviation from the ratio-scale is a function of changes in discrimination with changes in magnitude.[25] The *S*s construct units of their "uncertainty," not units reflecting ratios on a magnitude-scale. Torgerson also notes that this "uncertainty" is based on "differences," not "ratios." We specify just noticeable gains in stimuli and report on precision to plus or minus so many units. In contrast to a scale based on ratio-estimation, judgments of extreme stimuli reflecting the top of the scale are no rounder than those at the bottom of the scale. Hence, the use of numbers to specify a ratio is not an issue. This interpretation was supported by the evidence presented here in which for all the *S*s discrimination became poorer with increases in length or weight. The normal and sophisticated *S*s appeared a little better at noticing differences than the retarded *S*s, but still, all groups manifested the typical logarithmic type of psychophysical function characteristic of reporting differences instead of ratios.

The results of an analysis of variance of the data of pair-comparison showed the retarded *S*s were slightly poorer in their discrimination than the other *S*s. This is consistent with Holowinsky's finding that severely retarded *S*s did significantly worse in discriminating lengths of lines than normal *S*s.[26]

---

[23] Eisler, *op. cit.*, 243-253 (a); Torgerson, *op. cit.*, 1960, 21-31.
[24] Eisler, *op. cit.*, 243-253 (a); Torgerson, *op. cit.*, 1960, 21-31.
[25] Stevens, *op. cit.*, 1962, 29-39.
[26] I. Z. Holowinsky, Length discrimination as a function of IQ and motivation, *Trng. Sch. Bull.*, 61, 1964, 116-119.

A difference score between the true ratio and specified ratio of points was used to measure accuracy of the constant-sum judgments. An analysis of variance of these scores for the three groups showed dramatic differences. It was clearly evident that the most accurate judgments were made by the sophisticate group, and least accurate judgments were given by the retarded group.

Based on the data from this study, it appears that the accuracy of ratio-judgments relative to the sophistication of the *S*s was a function of their ability to use numbers as proportions of a standard stimulus rather than a function of their ability to experience these differences in "heaviness" or "longness."

From these findings, it should also be apparent that ratio-estimation is not an unbiased magnitude-scale. Both discriminative and magnitude-scales have their own unique problems. The solution to the dilemma of their inconsistent findings would appear to lie in a theoretical framework in which the assessment of *S*'s bias is done in such a way that we can allocate different fractions of bias to the reporting procedure versus judgmental error, in much the same manner as Guilford has proposed by distinguishing between the response and judgment continua.[27] Eisler has made one such attempt.[28] Other attempts will undoubtedly improve this situation. It would seem that the answer does not lie in accepting either scale as giving an unbiased estimate of subjective magnitude.

Summary and conclusion. This investigation was undertaken to study the relationship between psychophysical functions based on a pair-comparison and a constant-sum scale for length and weight over three levels of the *S*s' sophistication and ability. Previous evidence has shown that discrimination (pair-comparison) and magnitude (constant-sum) scales relate differently to the stimulus-continuum and are nonlinearly related to each other. The question is, which of the two scales is biased in giving a "correct" estimate of a specified sensory process.

Judgments of length of lines and heaviness of weights constituted the two tasks used in this study. Three groups of *S*s were used: (1) Twenty mentally retarded *S*s institutionalized at Pacific State Hospital with *IQ*s between 45-70; (2) twenty normal *S*s with a high school education; and (3) fifteen sophisticated *S*s, eight with the Ph.D. degree, and seven with the M.A. degree; all were familiar with scaling.

Findings. (1) The psychophysical function of the pair-comparison scale was negatively accelerated (concave downward) over length and weight for the retarded, normal, and sophisticate groups. (2) The psychophysical function of the constant-sum scale was positively accelerated (concave upward) over length and weight for the three groups. (3) A power function $\Psi = (S + K)^n$ described the relationship of the constant-sum scale to the magnitude of length and weight for all groups. The exponent "n" for both length and weight varied systematically with the *S*s' sophistication, *e.g.* the exponent was largest for the retarded *S*s, $n > 1.90$; next largest for normals, $1.69 > n > 1.60$; and least for the sophisticate group, $n < 1.40$. The magnitude of the exponents describes the degree to which these *S*s overestimated the relative value of the larger stimuli. (4) The accuracy of judgments for the pair-comparison data and constant-sum estimates was investigated separately. In general, the retarded *S*s exhibited poorer discrimination (particularly with respect to length) than the other *S*s, although the difference between the three groups was quite small in comparison to wide discrepancies found for the constant-sum estimates. In the latter estimates, the most accurate judgments were made by the sophisticate group and least accurate responses were given by the retarded (institutionalized) group.

---

[27] Guilford, *op. cit.*, 1954, 319-320.

[28] Eisler, The connection between magnitude and discrimination scales and direct and indirect scaling methods, *Psychometrika*, 30, 1965, 271-289.

Conclusion. Ratio-scales are most useful regarding quantitative relations among the stimuli. The pair-comparison scale is based on measures of uncertainty or differences between stimuli. In ratio-judgment, the numbers given increase exponentially as we use them to describe greater differences in the magnitude of two stimuli. In contrast to a scale derived from ratio-estimates, discriminative responses are not based on the use of numbers. Discrimination merely becomes poorer with stimulus-magnitude.

In ratio-judgments, overestimating increases with larger stimuli and does so as a function of *S*'s ability to use numbers.

Finally, the power-law applies to ratio-estimation and a negatively accelerative function (logarithmic) characterizes the psychophysical relation of the pair-comparison scale, regardless of the *S*'s education and intelligence. It would also appear that the pair-comparison scale has limited application to the judgment of length and weight because of unstable scale-values and that the constant-sum scale is definitely affected by a result of the *S*'s knowledge and experience in manipulating numbers.

MEASURING AMBIGUITY

By GERALD H. FISHER, University of Newcastle upon Tyne, England

The fundamental process involved in perceptual organization is that in which certain parts of a stimulus-pattern become distinguished from others in such a way as to be figural while the remainder of the pattern is relegated to the background. Figural aspects of stimuli differ from their context in that they have form, structure and contour. In comparison the ground is without form, being unstructured and relatively lacking in contour. Thus figure, in contrast with ground, is dominant and demanding of attention and the figural quality of part of a stimulus-array renders it available for subsequent, more definitive classification. Certain stimulus-configurations are ambiguous, however, in the sense that they may be perceived as representing more than one subject. Perhaps the most well-known examples of figures of this kind are those depicting 'my wife and my mother-in-law' prepared by Hill,[1] which was drawn to the attention of psychologists by Boring,[2] and the 'chalice and faces' figure first described by Rubin.[3] These patterns, along with such reversible figures as the Necker cube,[4] the Schröder staircase[5] and Sinsteden's windmill,[6] were employed originally in investigations designed to consider the mechanisms underlying spontaneous fluctuations in attention. Subsequently, their use has been extended to experimental and clinical studies in which hypotheses concerned with the influence of age, sex, motivation, set, along with cognitive, personality, and developmental variables upon features of perceptual organisation have been considered.

An important assumption implied in extension of the application of these materials is that the probability of perceiving each of the two aspects embodied within

---

* Received for publication June, 29, 1967. The research reported in this paper was supported by Grant No. 70/GEN/9617 from the Ministry of Defence.

[1] W. E. Hill, My wife and my mother-in-law, *Puck,* November 6, 1915.

[2] E. G. Boring, A new ambiguous figure, this JOURNAL, 42, 1930, 444-445.

[3] E. J. Rubin, *Synsoplevede Figurer: Studier i psykologisk Analyse,* 1915, 1-228.

[4] L. A. Necker, Observations on some remarkable phenomena seen in Switzerland; and an optical phenomenon which occurs on viewing of a crystal or geometrical solid, *Phil. Mag, 3 ser.,* 1, 1832, 329-343.

[5] H. Schröder, Über eine optische Inversion bei Betrachtung verkehrter, durch optische Vorrichtung entworfener, physischer Bilder, *Ann. Phys. Chem.,* 181, 1858, 298-311.

[6] W. J. Sinsteden, Über ein neues pseudoskopisches Bewegungsphänomen, *Ann. Phys. Chem.,* 187, 1860, 336-339.

them is the same. Should the appearance of each alternative and mutually exclusive aspect of ambiguous figures be other than equally probable, difficulties are raised both in submission of relevant numerical data to analysis and in generalising from the results of studies in which they have been used. When describing the 'wife and mother-in-law' figure, Boring observed that it ". . . *seems to be the best of the puzzle-pictures in the sense that neither figure is favoured over the other.*"[7] Since the aspect of the figure depicting the younger woman's face was observed by Leeper to appear upon 65% of occasions[8] and since Botwinick, Robbin and Brinley reported this aspect to be seen variously upon between 78% and 94% of occasions,[9] the assumption implied in Boring's statement appears to be questionable. Appreciating this problem, however, Botwinick prepared three versions of a 'husband and father-in-law' figure.[10] In one of these (C), the 'father-in-law' was emphasized, in (B) the 'husband', while in (A) an attempt was made to produce a figure in which the two aspects were equally dominant. Submitting the figures so devised to appropriate experimental consideration, Botwinick observed that 24 members of a sample of 51 *Ss* reported seeing the 'husband' upon first exposure of this latter figure. Accordingly, he considered it reasonable to assume that the probability of appearance of each feature of this figure was approximately the same.

EXPERIMENT I

Method and procedure. Five well-known and widely-used ambiguous figures were selected for study. These included the 'wife and mother-in-law', the 'husband and father-in-law', an example of a 'chalice and faces' figure, a version of Jastrow's 'duck and rabbit' figure[11] and the 'pirate and rabbit' scene prepared by Leeper.[12] In addition to these, three new figures were devised, the alternative subjects of which were a 'lion and a rose', a 'snail and an elephant' and a 'gypsy and a girl'. Taking each of these 8 figures as a model, 20 further drawings were prepared in such a way that, in 10 of them one feature was emphasised increasingly, the other aspect being emphasised similarly in the remaining 10. In this way 168 individual line-drawings were made. Twenty *Ss*, sophisticated in the sense of having been made fully aware of the purpose of the study, were asked to select; first, a central member of each set of 21 drawings in which both aspects appeared to be equally dominant; and secondly, 14 further figures, 7 on either side of the one first selected, graded in such a way that the difference between each adjacent pair was approximately the same. In the case of certain of these original materials, it should be noted that the *Ss* failed to agree upon acceptable figures to be included, whereupon further modifications were made to them until agreement was reached. Using this procedure 8 sets of materials were obtained each containing 15 figures, the dominant aspects of which varied between their 2 possible extremes. Seven of these 8 sets of drawings have been illustrated by Fisher,[13] the remaining set depicting the *'Gypsy and Girl'* being illustrated in Fig. 1.

---

[7] Boring, *op. cit.*, 445.

[8] R. W. Leeper, A study of a neglected portion of the field of learning—The development of sensory organization, *J. genet. Psychol.*, 46, 1935, 41-75; esp. 62.

[9] Jack Botwinick, J. S. Robbin, and J. F. Brinley, Reorganization of perceptions with age, *J. Geront.*, 14, 1959, 85-88.

[10] Botwinick, Husband and father-in-law—A reversible figure, this JOURNAL, 74, 1961, 312-313.

[11] Joseph Jastrow, *Fact and Fable in Psychology*, 1900.

[12] Leeper, *op. cit.*, 62.

Each individual member of this total set of 120 drawings was photographed and prepared in the form of a photographic transparency. Mounted in slides, these were presented to seven independent groups of naïve *Ss*, a total of 200 people, of whom 131 members were men, ranging in age from 16 to 72 yr. The individuals constituting the experimental sample were drawn from widely differing walks of life such that they may reasonably be considered representative of a cross-section of the populace of the English countries of Northumberland and Durham. The figures were presented to the *Ss* in a randomly determined order, this being changed for each sample participating in the experiment.

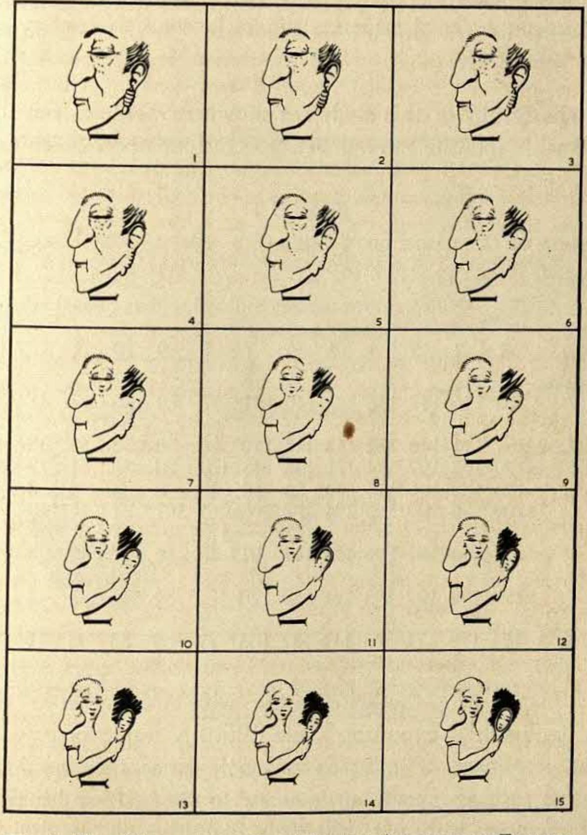


FIG. 1. THE 'GYPSY AND GIRL' SET OF AMBIGUOUS FIGURES

---

[13] G. H. Fisher, Materials for experimental studies of ambiguous and embedded figures, *Res. Bull. No. 4,* 1966, Dept. of Psychol., University of Newcastle upon Tyne.

Instructions. All the *S*s were given instructions verbally in the following form:

> On this screen [screen indicated] I will show you a number of line-drawings which depict various human and animal subjects. Please look at each of these carefully and indicate what you can see by writing an appropriate descriptive word, or phrase, on the piece of paper I have just given to you. Do this immediately you identify the subject of each drawing. Should you see more than one thing, or if what you see changes in some way, please be sure to indicate the first aspect you see.

Each slide was exposed for a period of 15 sec., 2 sec. intervening between successive presentations. A period for rest, 2 min. in duration, was made after presenting each group of 30 slides.

Results. The results of this study are shown in Table I. The figures entered in this table indicate the number of occasions upon which each of the aspects first named (*e.g.* the 'wife' in the 'wife and mother-in-law' figure) were seen upon initial exposure. Little difficulty was experienced in identifying the responses of most *S*s uniquely since they quickly realized that responses such as '*female*' or '*woman*' to the 'wife and mother-in-law' figure, for example, were not sufficiently definitive and modified them accordingly.[14] It was necessary to discard 14 protocols in which distinctions of this kind failed to be made.

TABLE I

THE NUMBER OF OCCASIONS UPON WHICH ONE ASPECT OF THE EIGHT SETS OF AMBIGUOUS FIGURES WAS SEEN UPON FIRST EXPOSURE

Figure	Number of responses indicating first-named aspect seen														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Wife and Mother-In-Law	160	146	147	146	143	140	101*	97	89	66	62	41	29	16	21
Husband and Father-In-Law	197	171	163	152	118	111	110*	86	78	50	58	38	18	6	7
Chalice and Faces	190	192	192	194	171	138	143	120	80	58	61	72	60	49	54
Rabbit and Duck	195	188	190	133	89*	56	53	59	8	8	4	2	3	3	0
Lion and Rose	200	196	200	198	194	192	168	161	107*	27	16	9	3	1	0
Snail and Elephant	197	184	192	196	193	117	91*	40	16	17	6	3	3	3	5
Pirate and Rabbit	192	199	192	193	194	180	120	106*	59	46	33	2	8	2	5
Gypsy and Girl with Mirror	171	178	171	163	144	132	103*	79	68	53	43	18	11	10	7

---

[14] The two aspects of the set of ambiguous figures illustrated in Fig. 1 were intended to represent a gypsy and a girl holding a hand-mirror. Although the two aspects of this figure were obviously distinguished in the *S*s' responses, a number of different responses were made by way of description of each. To the male figure, in addition to the response 'Gypsy,' the descriptions 'dowager' and 'dowager Duchess,' 'Spaniard,' 'Mexican,' 'gigolo,' 'pirate' and 'man with big nose' were given. All the *S*s described the female figure as being female, usually as 'young' and frequently as 'blonde.' Interestingly, a significant proportion of female sub-

The standard deviation of a sampling distribution of two equally probable and mutually exclusive responses made independently upon 200 occasions is $(200 \times 0.5 \times 0.5)^{1/2} = 7.1$, its mean value being $200 \times 0.5 = 100$. It may be considered justifiable to accept a figure as being ambiguous, in the sense that the appearance of each of its two alternative aspects is equally probable, if the number of responses indicating one of them to become apparent falls within the range plus or minus two standard deviations about the mean of the sampling distribution. Accordingly, the acceptable range for achievement of the condition of equivocality is 86-114 responses. At least one member of seven of the eight sets of materials studied proved acceptable in these terms. The figure to which the number of responses made is closest to the mean of the ideal sampling distribution is indicated in the table. Six of these equi-probable figures are illustrated in Fig. 2, the seventh is the seventh member of the set shown in Fig. 1.

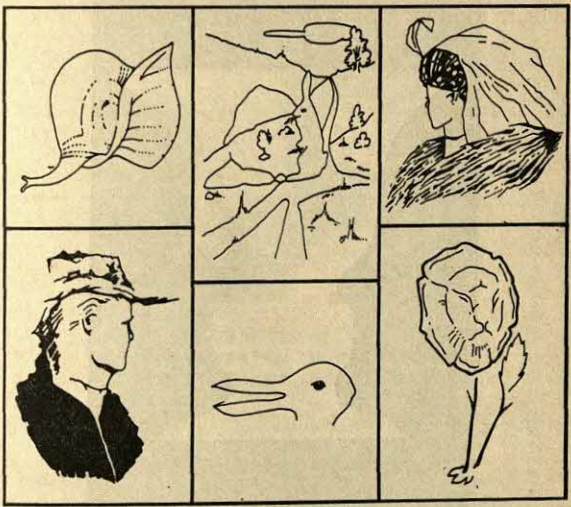


Fig. 2. Six Ambiguous Figures the Probability of Perceiving Either Aspect of Which Was Observed to be Approximately Equal

In the case of the set of figures depicting either 'Chalice and Faces' the attempt to produce a figure, the probability of perceiving either aspect of which is the same, failed. To the figure numbered 8 in the series, 120 responses were made indicating the 'chalice' aspect to be seen, while the adjacent figure, numbered 9, received 80 similar responses. Each of these values falls at a point 2.82 standard deviations from the mean of the ideal distribution. The probability of this value representing an equivocal stimulus is less than 0.01, on the appropriate null-hypothesis, considered against relevant degrees of freedom. Hence, neither of these figures can reasonably

---

jects identified this aspect as representing 'a female holding a baby'. It is likely that further differences of this kind will become apparent if these figures are used with *S*s from different cultures.

be accepted as being equally probable in the appearance of their two alternative aspects. A 'chalice and faces' figure modified in such a way as to fall between the 60% and 40% points represented by Figs. 8 and 9 in this series is illustrated in Fig. 3. While it may be reasonable to consider the chalice and the faces to appear with equal frequencies, it should be noted that this figure remains to be submitted to experimental consideration.

EXPERIMENT II

Ambiguous stimulus-patterns, the figural aspects of which reverse spontaneously, have provided the subjects of many experimental inquiries. While it has been shown variously that a number of peripheral factors including eye-movements, points of fixation, and contrast influence fluctuation in the perceived aspects of such figures, it is generally agreed that the principal mechanism underlying these changes is centrally organised. Many suggestions have been made as to the possible mode of operation of such a mechanism. While none of these is highly definitive, in general terms, they all imply the existence of excitatory and inhibitory, biochemical states, the balance of which changes as a function of processes operating in centrally organized neurological systems. A variety of explanations offered by Gestalt psychologists (*e.g.* Köhler) are expressed in such terms as these.[15] Further, attempts are being made increasingly to consider relationships between figure-ground fluctuations and

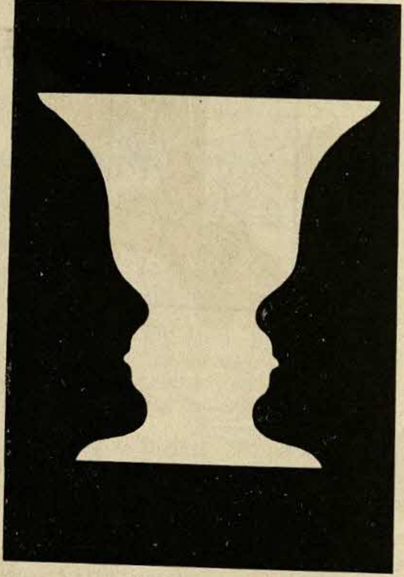


FIG. 3. A 'CHALICE AND FACES' FIGURE FALLING BETWEEN THE 40% AND 60% POINTS OF THE EXPERIMENTAL SERIES

personality variables thought to be related to states of 'excitation-inhibition' imbalance.

---

[15] Wolfgang Köhler, *Dynamics in Psychology*, 1940, 1-156.

Many criteria may be invoked to consider the general problem of the locus of organization of perceptual mechanisms. One of these, perhaps somewhat neglected by psychologists, is to consider the characteristic time-constants of the particular coding system involved. Most reports of the sequential time-course of fluctuations in the perception of ambiguous figures indicate a periodicity of the order of a few seconds in duration. The mean values of a number of available reports, including the extensive studies carried out by Porter and Mori, fall in the range approximately 3-5 sec.[16] This has been interpreted as evidence indicating a central locus of organization. In view of these considerations, the interesting problem arises of the possibility of inducing changes in the perception of ambiguous figures which are other than spontaneous. If such changes could be induced, the implication would be that the periodicity of fluctuations of this kind is capable of being manipulated.

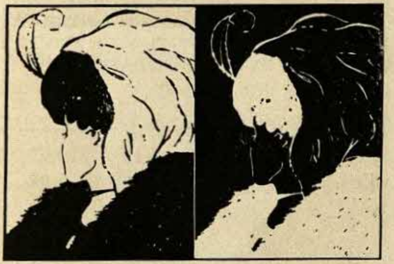


FIG. 4. THE ORIGINAL 'WIFE AND MOTHER-IN-LAW' FIGURE AND ITS 'REVERSED-NEGATIVE'

Method and procedure. Fluctuations in the perception of aspects of ambiguous figures may be induced readily using stimuli of the kind illustrated in Fig. 1. A figure, the probability of perceiving one of the embodied aspects of which is high, could be presented alternately against a second figure, the probability of perceiving the same aspect of which is correspondingly low. This method would not be entirely satisfactory, however, since the ranges over which these stimuli have been drawn are such that their configurations differ appreciably both in contour and in form. It is possible to control all aspects of form and contour, however, if each figure is presented alternately against what may be described as its *'reversed-negative'*. An example of such a stimulus is shown in Fig. 4. The left-hand figure is the original 'wife and mother-in-law' figure. This figure is a 'positive' (*i.e.* black on white). The right-hand figure has been prepared in such a way as to represent an exact 'negative' version (*i.e.* white on black) of the original figure, this being reversed left to right. Thus, the 'reversed-negative' remains in the same spatial orientation as the 'positive' figure, the two figures being otherwise identical, both in contour and in form. Figures of this kind may be prepared readily using a simple photographic procedure.

---

[16] E. L. H. Porter, Factors in the fluctuation of fifteen ambiguous phenomena, *Psychol. Record,* 2, 1938, 231-253; G. F. Mori, On the fluctuation of ambiguous figures, *Atti. Fond. G. Ronchi,* 18, 1963, 308-322.

The equi-probable 'Wife and Mother-in-Law' figure, illustrated in Fig. 2, was prepared both in the form of a positive and a reversed-negative. The figures were shown individually to a sample of 200 *S*s. These *S*s were not entirely naïve since they had already participated in Experiment I. Nevertheless, they had no knowledge of the purpose of this particular part of the study. The stimuli were presented in a two-channel tachistoscope, each *S* being shown first either the 'positive', or the 'reversed-negative', and asked to 'hold' one aspect, or the other, for a period of 10 sec., whereupon the figure was replaced immediately either by its complementary partner, or by the same figure. The stimuli first presented, the aspects to be 'held' and the figures presented upon replacement were selected at random among the *S*s, who were asked to indicate which aspect they saw when the second figure was presented.

TABLE II

FREQUENCIES OF APPEARANCE OF EITHER ASPECT OF THE 'WIFE AND MOTHER-IN-LAW' FIGURE SUBSEQUENT TO PRESENTATION OF EITHER ITS 'POSITIVE' OR ITS 'REVERSED-NEGATIVE'

Prefixation	Aspect 'held'	Stimulus	Aspect seen	
			Wife	Mother-in-law
positive	wife	positive	92	108
		negative	35	165
	mother-in-law	positive	117	83
		negative	147	53
negative	wife	positive	51	149
		negative	90	110
	mother-in-law	positive	161	39
		negative	121	79

Results. The results of this study are entered in Table II. For clarity they are shown in such a way that the aspects to be 'held,' the figures replaced and the aspects apparent upon replacement may be seen readily. In the case of each of the eight sequences of presentation, there is a tendency for the aspect of each figure seen to change upon replacement, irrespective of whether the stimulus represents a positive, or reversed-negative, version of the prefixation figure. There is, however, a clear difference between replacement of a figure with the same figure and with a complementary version of it. The probability of occurrence by chance of the data entered in the last two columns of the table may again be estimated by considering them against an ideal sampling distribution having a mean value of 100 and a standard deviation of 7.1. All complementary sequences fall well outside the range 86-114 responses, two standard deviations on either side of the mean, while only one of the control sequences does so. These results may thus be interpreted as providing a significant indication that changes in the aspects of ambiguous figures seen may be induced by manipulating their figure-ground characteristics in this way. There were no sex differences in the results and no apparent differences in the facility with which men and women were able to 'hold' the figures in either of their aspects. A number of brief studies, using the other figures illustrated in Fig. 2, and with Necker cube figures prepared in a similar way, indicate that fluctuations of this kind may be induced in them also using the same technique.

EXPERIMENT III

The experiments reported here are concerned with the 'figure-ground' phenomenon. This phenomenon may be observed readily in such figures as the 'wife and mother-in-law' in which one sees either the 'wife' or the 'mother-in-law' but never both at the same time and never anything else. Describing the characteristic nature of this phenomenon Vernon comments ". . . *Only the figural part of the field is central in awareness and is clearly perceived at any one time.*"[17] Vernon, and others, attribute this finding to the Gestalt school of psychologists, in particular to Rubin, describing it as *'the figure-ground hypothesis'*.[18] There appears in our view, however, to be a sense in which the figure-ground hypothesis remains to be tested satisfactorily. Early investigations, inspired by the classical philosophical question of whether the mind could apprehend more than one object at any one time, have been repeated so frequently as not to require elaboration here (see Woodworth; and Popper).[19] These studies are described under the general headings of 'span of attention' and 'span of apprehension.' Hunter and Sigler asked the question *"How many objects can be grasped in a single moment of consciousness?,"*[20] and attempts to answer this have provided the empirical foundations of information theory.[21] Experiments of this kind, however, have usually been carried out with abstract, or otherwise meaningless, stimuli such as dot and form patterns presented against homogeneous backgrounds. As far as we are aware the related question, can equivocal stimuli of the kind characterised by ambiguous figures be seen to represent *more than one object simultaneously*, remains to be considered.

---

[17] M. D. Vernon, *A Further Study of Visual Perception*, 1952, 45.

[18] Vernon, *op. cit.*, 41.

[19] R. S. Woodworth, *'Experimental Psychology,'* 1938, 684-704. K. R. Popper, Philosophy of science: A personal report, in C. A. Mace (ed.), *British Philosophy in the Mid-Century*, 1957, 175-176.

[20] W. S. Hunter and Marian Sigler, The span of visual discrimination as a function of time and intensity of stimulation, *J. exp. Psychol.*, 26, 1940, 160-179. For a bibliography and brief history of the experiments on the "Range of Attention", see N. F. Gill and K. M. Dallenbach, this JOURNAL, 37, 1926, 247-256; also, A. D. Glanville and K. M. Dallenbach, this JOURNAL, 41, 1929, 207-236.

[21] G. A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information, *Psychol. Rev.*, 63, 1956, 81-97.

Method and procedure. A good deal of time was spent in the attempt to devise materials suitable for considering the hypothesis of 'unity of perceptual span.' Early examples were inspired by the works of Escher, the Dutch artist,[22] and Vigeland, the Norwegian artist and sculptor. Some of these are illustrated by Fisher.[23] The drawback of these materials was that it was found difficult to overcome the tendency to reproduce figures which, although having common peripheral outlines and profiles, extended spatially in such a way as to be largely independent of each other. Hence, they occupied different parts of the visual field. These drawings, although resembling the 'Chalice and Faces' figure, differed from 'Wife and Mother-in-Law' type figures in which each embodied aspect is largely contained within the other, having common spatial extensions and occupying the same parts of the visual field. Having succeeded in overcoming this tendency, however, materials of the kind illustrated in Fig. 5 were devised. These figures have three, five, and seven aspects. Each contains a full-face figure. In addition to this the Italian, Chinese, Red Indian, and Female figures contain two inward-looking profiles. The Jester, Old Woman, Arab, and Egyptian figures each contain two further aspects, these being obtained by the addition of alternative treatments to the jawlines of the two facing profiles. Embodied in the Judge, Clown, Toby-jug, and Gangster figures are seven aspects; two outward-looking faces having been added to five-aspect figures.

In a pilot-study conducted with these materials, 40 *Ss* were asked to indicate what they could see when first shown these figures. Their reports were recorded on magnetic tape. These observations were extremely difficult both to evaluate numerically and to interpret. The problem was to determine if one or more aspects of each figure could be seen *simultaneously*. To achieve this, it was necessary to ask the *Ss* questions, some of which it was difficult to avoid casting in a leading form. Five independent *Es*, fully conversant with the purpose of the study, found themselves unable to reach agreement upon the import of the responses of 18 members of the pilot-sample. Accordingly, an attempt was made to carry out the study in a somewhat more objective way.

Each member of a sample of 100 unsophisticated *Ss* was shown one example of each of the three-, five-, and seven-aspect figures. These were selected from the set illustrated in Fig. 5 and they were presented in a randomly determined order for number of aspects among the *Ss*. The first question asked was always: "*Please tell me what you can see in this drawing.*" Almost without exception, the response to this question indicated that the first feature of the figure to become apparent was that representing its full-face aspect. If no further observation had been made spontaneously within a period of approximately 15 sec., *Ss* were asked: "*Can you see anything else?*" This question was repeated in the case of the five- and seven-aspect figures until the *Ss* reported being unable to see any further aspects. This procedure succeeded in eliciting responses indicating all available aspects to be seen by 58 *Ss*. The features most resistant to observation were those defined by the alternative jawline treatments, the outward-looking faces in the seven-aspect figures being seen readily. When all aspects of the figures failed to become apparent spontaneously, the remaining features were pointed out to the *Ss* by indicating the contours of the profiles with the point of a pencil. Subsequent to this procedure, all but 2 *Ss* reported being able to determine all aspects of each of the figures with which they were presented.

---

[22] M. C. Escher, '*The Graphic Works of M. C. Escher*,' 1960, 10-27.
[23] Fisher, *op. cit.*, *Res. Bull. No. 4*, 1966, 10.

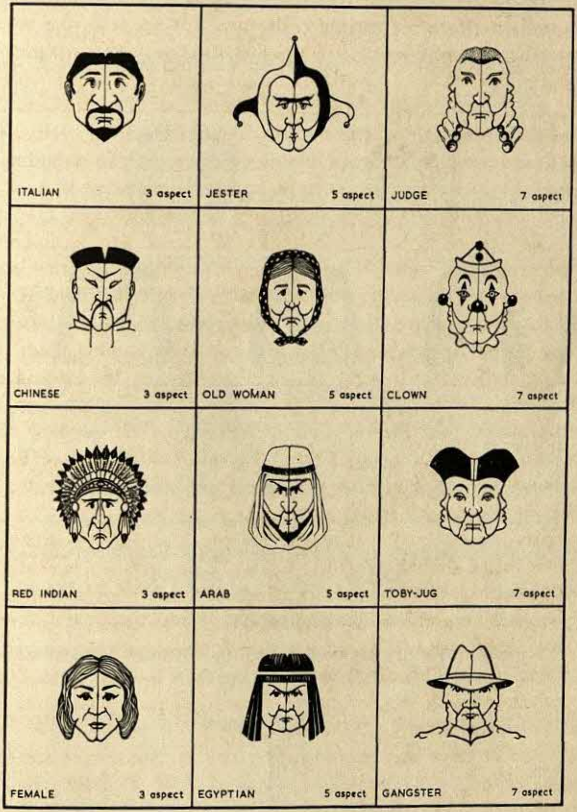


FIG. 5. THREE-, FIVE-, AND SEVEN-ASPECT AMBIGUOUS FIGURES

Materials of the kind illustrated in Fig. 6 were then shown to *Ss*, all of whom, except the two members mentioned above, agreed readily that they represented the aspects of the figures which they were able to see. The *Ss* were now asked the further question: *"Can you see both* [*or more than one*], *features of these figures at the same time? By this I mean (example) can you see the full-face of the clown and the inward (or outward) looking faces simultaneously?"*

Results. The results of this experiment are entered in Table III. They refer to the number of *S*s who reported either one, or more, aspects of each figure to be seen without further qualification. Discussion with some *S*s who reported being able to see more than one alternative at the same time revealed that, had the instructions been more definitive, they would have responded differently, nevertheless, the data referring to these responses were preserved in the appropriate cells.

These data require little by way of statistical consideration. They appear to indicate clearly that, for the majority of the *S*s, even when more than the two alternatives of the classical ambiguous figure are available, one figure or aspect only becomes figural at any given time. Naturally, the simultaneous figural quality of exactly symmetrical features of these figures is excepted from this statement.

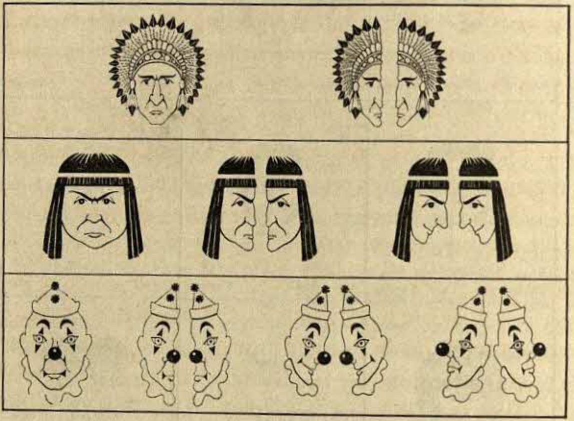


FIG. 6. EXAMPLES OF NON-AMBIGUOUS FORMS OF MULTI-ASPECT AMBIGUOUS FIGURES

TABLE III

THE NUMBER OF ASPECTS OF AMBIGUOUS FIGURES REPORTED TO BE SEEN SIMULTANEOUSLY

Aspects included	One aspect seen	More than one aspect seen
3	93	7
5	91	9
7	83	17

Discussion and conclusions. The experiments described in this paper have been conducted both with well-known figures and with spatial patterns devised especially for the study. The aspects which these figures assume are ambiguous, in the sense that they may be seen to represent more than one subject. Figures of this kind have been used in a wide variety of experimental studies in some of which certain implications and requirements may have been overlooked. The principal difficulty is the assumption that the probability of appearance of either aspect of an ambiguous spatial pattern is the same. This condition must be achieved if interpretation of the results of investigations in which they have been used are to be of widest generality. The frequencies of appearance of either aspect of a set of ambiguous figures, individual members of which vary between two extremes, according to the procedure devised for preparing the materials used in this experiment, may be understood as defining a sampling distribution of mutually exclusive responses differing in the probabilities with which they are made. If p is the probability of one response being given and q the probability of the alternative, then the probabilities of all possible combinations of responses are given by the coefficients of the binomial series generated by expansion of the expression $(p + q)^n$. When $p = q = 0.5$ the distribution of binomial coefficients approaches that of the Gaussian, or normal distribution, function with increasing 'n'. The mean value of this distribution is $n \times p$ and its standard deviation $(n \times p \times q)^{1/2}$. Bearing this in mind, it is possible to estimate the probability of any given combination of responses being made to spatial patterns of the kind under consideration. Here special attention has been paid to the mean value of the distribution of binomial coefficients represented by figures which are truly ambiguous in the sense of each alternative response to them being made with equal frequency. Use of figures which fail to meet this criterion yields distributions of responses which in a statistical sense are skewed. It is possible that departure from essential normality of such distributions may be such as to severely restrict the generality of interpretation of experimental investigations in which figures of unequal ambiguity have been used. It appears reasonable to claim that the condition of equivocality has been achieved in the case of seven of the eight sets of figures prepared for the present study.

In his *Outline of Abnormal Psychology* McDougall proposed that the rate of spontaneous reversals of ambiguous-figures indicates the speed with which neural impulses travel in the central nervous system.[24] The speed of such impulses, he suggested, is related to extraversion-introversion, being slow for extraverts and fast for introverts. Cameron, using the Schröder staircase figure, observed that fluctuation rates for manic-depressive patients were slower than those for normals.[25] Eysenck reported that "*the reversal rate* (for the Rubin 'Chalice and Faces' figure and for the Maltese cross, which appears to be a version of Rubin's reversible-sectors) *is rather*

---

[24] William McDougall, *Outline of Abnormal Psychology*, 1926.
[25] D. E. Cameron, Studies in depression, *J. ment. Sci.*, 82, 1936, 148-161.

lower for neurotics than for normals" and that female subjects show both higher rates of reversal and a higher degree of control over reversals than males.[26] Jaensch investigated reversals of perspective in relation to personality types,[27] while an attempt to evaluate spontaneous figure-ground reversals in a clinical test situation and to use this as a measure of 'rigidity' was described by Frenkel-Brunswik.[28] Findings of this kind, along with the vast number of studies on so-called 'satiation' processes, have been interpreted as pointing directly toward a locus of figure-ground organization and reorganization at a central level.

Detailed discussion of the physiology of the eye and the system of visual projection is out of place in the present context. Nevertheless, it is well-known that the functional elements of the retina differ in their modes of operation and that these differences ramify with differences in neural transmission toward the chiasma and beyond. The most obvious avenue for development of an explanation of the facility with which reversals may be induced in ambiguous figures, prepared in such a way that their figure-ground characteristics are exactly complementary, lies in consideration of 'on,' 'off,' and 'change-over' responses to variations in light stimulation. Both Kuffler and Granit demonstrated the antagonistic and mutually exclusive nature of these responses at retinal level, these presumably serving to enhance contrast and facilitate the formation of contours.[29] From the complementary nature of these retinal processes may be developed a persuasive theory of figure-ground reversals at a peripheral level. Recording from 'binocular' units in the striate cortex of the cat, however, Hubel and Wiesel observed interaction between excitatory and inhibitory states in the two homologous receptor fields.[30] The nature of this interaction is such as to conceivably confound the argument in favor of a peripheral level of organization. Whether repetition of the present experiment in monocular rather than binocular vision would distinguish between possible loci of organization of the phenomenon under observation is a matter for speculation, which is probably premature without further evidence of the structural organization and functional repertoire of the visual projection system. Nevertheless, it may be claimed that complementary modes of perceptual organization of the binocular visual field have been demonstrated here and that the nature of these may be such as to lead toward further understanding of the mechanisms underlying figure-ground reversals.

Insofar as the 'figure-ground phenomenon' implies an hypothesis concerned with the possibility of the visual projection-system being in more than one state of organization simultaneously we would suggest that, hitherto, it has not been tested experimentally. Commenting upon the artistic treatment of Dali's painting 'Slave Market with Disappearing Bust of Voltaire,' detail of which is shown in Fig. 7, Carraher and Thurstone write

---

[26] H. J. Eysenck, *Dimensions of Personality*, 1947, 108.

[27] E. R. Jaensch, *et al.*, *Studien zur Psychologie menschilder Typen*, 1930

[28] E. Frenkel-Brunswik, Intolerance of ambiguity as an emotional and perceptual personality variable, *J. Person.*, 18, 1949, 108-143.

[29] S. W. Kuffler, Discharge patterns and the functional organization of the mammalian retina, *J. Neurophysiol.*, 16, 1953, 37-68; Ragnar Granit, *Receptors and Sensory Perception*, 1955, 1-367.

[30] D. H. Hubel and T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.*, 160, 1962, 106-154.

"the head of Voltaire and the two coifed nuns have been subtly altered so that the perception of one depends upon the presence of the other. Although the figures are embedded within the same contour, it is difficult to experience them simultaneously."[31] This observation underlines the problem on hand. Can more than a single aspect of ambiguous figures be seen *at one and the same time?*

Three problems arise from these considerations. First, is the question meaningful? If more than one aspect of an ambiguous figure can be seen simultaneously the implication appears to be that the visual projection-system can be in two different functional, and perhaps structural, states at one and the same time. Logically this seems unlikely; nevertheless, inasmuch as the organization of figure and ground is described as an hypothesis relevant to the appearance of spatial patterns, the possibility of differential states of organization being present simultaneously cannot be denied axiomatically. Wittgenstein drew attention to the logical connotations of what he described as *'noticing an aspect'* of an ambiguous spatial pattern but appeared to misunderstand the psychological implications of fluctuation in figure-ground organization when denying that *"'Seeing as' . . . is not part of perception."*[32] Our view is that consideration of what a particular figure is seen *as* and how it is seen *as such* constitutes a very large part of the legitimate subject matter of perception.

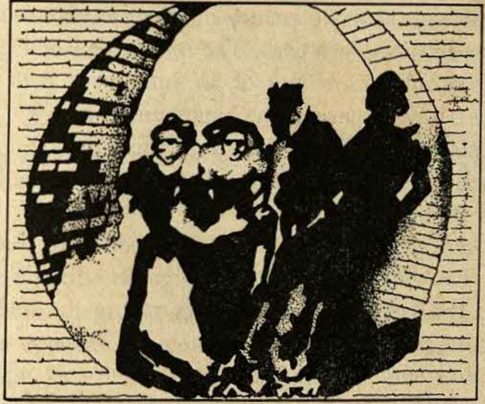


FIG. 7. DETAIL AFTER 'SLAVE MARKET WITH DISAPPEARING BUST OF VOLTAIRE'
BY SALVADOR DALI

---

[31] R. G. Carraher and J. B. Thurstone, *Optical Illusions and the Visual Arts,* 1966, 96.

[32] Ludwig Wittgenstein, *Philosophical Investigations,* 1953, 193-216.

Secondly, if we are in fact concerned with the empirical testing of an hypothesis, what phenomenon is under observation? This question appears to be gratuitous since it seems obvious that the phenomenon should be described as that of 'figure and ground.' But the figure-ground hypothesis allows for two categories, or alternatives, only. If more than two alternatives were available, the topic of discussion might reasonably be described as 'the *span of perception.*' This form of words provides one of the subheadings of Graham's commentary upon visual perception. In this, however, he points out the "*. . . need for an understanding of behaviour processes that depend on different classes of discriminative stimuli . . .*",[33] thus emphasising the necessity for studying more structured and meaningful stimulus-patterns than it is customary to employ in experiments concerned with the organization of perceptual processes. The implication of the figure-ground hypothesis is that only one aspect of an ambiguous pattern can become figural at any one time. Hence, if the hypothesis is true it must follow that perception in this sense has unit-span. It is this implication of the figure-ground hypothesis which, in our view, has not previously been submitted to experimental enquiry.

Third, how can such an hypothesis be tested? The crucial question is whether more than one aspect of an ambiguous stimulus-pattern can become figural at the same time. But the figures usually used for consideration of the figure-ground hypothesis contain two aspects only. When observing the "Wife and Mother-In-Law" figure, for example, one sees *either* the 'wife' *or* the 'mother-in-law' but never both at the same time, and never anything else, since only these two alternatives are available and the two categories are mutually exclusive. Suppose, however, that one, or more, further alternatives were to be available. Would they be perceived sequentially, alternating in aspect in the manner characteristic of conventional ambiguous figures, or simultaneously, more than one aspect being perceived figurally at the same time, thus revealing perception to have span longer than unity? While there may be a variety of techniques for achieving a situation in which the figure-ground hypothesis may be tested satisfactorily, the restrictions which it is necessary to place upon the stimulus-materials used appear to allow of one method only. This is to design an experimental situation in which more than two features of an equivocal stimulus-array can become figural when placed in the visual field. An attempt has been made to test this hypothesis by preparing ambiguous figures with more than the conventional two alternative aspects, considering these against appropriate non-ambiguous figures. The results of this experiment indicate that insofar as 'the figure-ground hypothesis' can be considered empirically and if it can be claimed to have been tested here, then it is upheld. Accordingly, it seems reasonable to conclude that, in this context at least, perception has unit-span.[34]

---

[33] C. H. Graham, Visual perception, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951, 902.

SUMMARY

Three experiments are reported. In the first, seven ambiguous figures, the probabilities of perceiving either aspects of which are approximately the same, are described and illustrated. In the second, the possibility of inducing fluctuations in the appearance of alternative aspects of ambiguous figures which are other than spontaneous is considered. In the third, an attempt is made to submit the 'figure-ground hypothesis' to experimental test.

---

[34] Reproduction of the full sets of figures prepared for this study is clearly impracticable in a paper of this length. Copies of all the materials used in the experiments may be obtained from the author on request.

TORQUE SENSITIVITY AS A FUNCTION OF KNOB RADIUS AND LOAD

By BURRTON WOODRUFF and HARRY HELSON, Kansas State University

In a previous study Woodruff[1] and Helson[2] measured sensitivity to torque by requiring *S*s to lift a rod with attached weight by grasping one end of the rod.[3] When a rotatory component is added to the stimulus a new dimension of sensitivity enters which can be measured independently of classical weight-sensitivity. The torque component was varied by moving a weight different distances from the point where the rod was held. Thus, although the weight of the stimulus was constant, the torque exerted on the fingers varied with the distance of the weight from the end of the rod. In the experiment reported here, sensitivity to torque was measured by requiring *S*s to turn knobs of various sizes against various loads.

There are several reasons why the task of knob-turning was chosen over that of the lifted-rod in this investigations: first, knobs are commonly used as control devices in which a force must be exerted tangentially to the line of motion against various loads; secondly, some variables are more easily studied when knobs are used instead of rods or levers. With lifted-rods it is much less convenient to assess effects of variables in commonly used devices, such as radius, inertia, mechanical advantage, and load of manipulanda. Finally, knob-turning was employed in this study because it simulates not only the design of manipulanda but also brings into play muscles and motions more commonly found in everyday life than does the lifted-rod previously employed.

Surprisingly little work has been done to investigate the sensitivity of individuals in a situation where torque sensitivity is in question. Hoisington employed stimuli very similar to those used by Woodruff and Helson, but he was interested in the nonvisual perception of length rather than in sensitivity to torque.[4] The only studies that concern themselves with torque sensitivity deal with terminal thresholds. For example, Sharp studied the maximal amount of torque that could be exerted on a

---

* Received for publication June 5, 1967.

[1] NIMH pre-doctoral trainee, Grant No. MH 8359.

[2] Now at York University, Toronto, Canada.

[3] B. G. Woodruff and Harry Helson, Torque: A new dimension in tactile-kinesthetic sensitivity, this JOURNAL, 78, 1965, 271-277.

[4] L. B. Hoisington, On the non-visual perception of the length of lifted rods, this JOURNAL, 31, 1920, 114-146.

knob of given radius and surface design.[5] He found that as knob radius increased from 0.25 to 2.50 in. maximal exertable torque increased. He also verified the intuitively obvious fact that more torque can be exerted on knurled knobs than on smooth knobs.

This paper presents the results of research in which torque-sensitivity was measured as a function of knob-diameter and load through the use of knobs of various diameters and loads of various amounts.

Method: (1) Apparatus. The apparatus consisted of a 45-cm. steel rod (1.27-cm. diameter) mounted on two ball-bearing assemblies so that the rod could be rotated freely. Knobs were attached to the rod by means of a drill chuck mounted on one end of the rod. The *Ss* compared the series weights, hereafter called loads, suspended from the rod with the standard weight by turning the knob so as to raise the weight from a fixed position a specified amount.

TABLE I

STANDARD AND SERIES LOADS IN GRAMS USED WITH EACH KNOB-SIZE

Knob-radius (cm.)	Standard load (gm.)	Series loads (gm.)				
1.27	50	35.0	42.5	50.0	57.5	65.0
	100	65.0	82.5	100.0	117.5	135.0
	200	160.0	180.0	200.0	220.0	240.0
2.54	50	26.0	38.0	50.0	62.0	74.0
	100	65.0	82.5	100.0	117.5	135.0
	200	140.0	170.0	200.0	230.0	260.0
3.81	50	26.0	38.0	50.0	62.0	74.0
	100	60.0	80.0	100.0	120.0	140.0
	200	140.0	170.0	200.0	230.0	260.0

Round knobs were cut from 0.25 in. plywood, sanded, and painted flat black. The knobs were 1.27, 2.54, and 3.81 cm. in radius (1, 2, and 3 in. in diameter).

In a pilot study, stimulus-series of 5 loads were chosen for each of the standards (*Std.*) which were 50, 100, or 200 gm. The step-interval between adjacent series-stimuli was so adjusted that the lightest stimulus produced approximately 15% heavier and the heaviest series stimulus 85% heavier than standard judgments using the method of constant stimulus-differences. Because knob radius affected the proportion of heavier judgments, three different sets of loads were employed with each knob size, as shown in Table I.

(2) Subjects. Twelve *Ss* (6 men and 6 women) were recruited from classes in general psychology and paid for participation. The average age of the men was 18.0 yr. and of the women, 18.8 yr. All the *Ss* were righthanded. Each *S* participated in one 50-min. session a day for three days a week until the experiment was completed.

---

[5] E. D. Sharp, Maximum torque exertable on knobs of various sizes and rim surfaces, *Tech. Doc. Rep.* No. MRL-TOR-62-17. Behavioral Sciences Lab., Wright-Patterson AFB, March 1962.

(3) Design. The experiment was set up as a complete factorial arrangement of treatments giving 9 conditions (3 knob radii × 3 load series). The conditions were presented in a random order with all the *S*s receiving the same order. Because this procedure might lead to changes in threshold with practice, a spot check was made by rerunning the first condition after all the data had been collected. The following is the order in which the conditions were presented: 2.54 cm. knob with 100 gm. *Std.;* 1.27 cm. knob with 100 gm. *Std.;* 2.54 cm. knob with 200 gm. *Std.;* 3.81 cm. knob with 200 gm. *Std.;* 1.27 cm. knob with 50 gm. *Std.;* 2.54 cm. knob with 50 gm. *Std.;* 3.81 cm. knob with 50 gm. *Std.;* 1.27 cm. knob with 200 gm. *Std.;* 3.81 cm. knob with 100 gm. *Std.;* and finally, a repetition of the first condition—the 2.54-cm. knob with 100-gm. *Std.* No differences due to practice were found between the original and rerun observations with the first condition.

(4) Procedure. The method of constant stimulus-differences was used to measure difference thresholds under the 9 conditions. Each *S* made 150 judgments under each condition—30 judgments for each standard-variable combination. Both time-order presentations were used in a single session. One-half of the *S*s randomly received each time-order for the first 75 trials on any given day.

To minimize any effects of order of presentation, 9 different sequences were used. Each sequence was constructed by randomly choosing 4 of the 120 ways 5 stimuli can be ordered. The entire set of 150 trials for a session was derived by repeating these 4 randomly chosen orderings. Sequences were randomly assigned to the 9 experimental conditions so that all 9 sequences were used in each condition. The same sequence was not repeated with any *S*.

When *S* entered the experimental room he saw a large black plywood shield with the chuck protruding from it. The chuck was 94 cm. from the floor. The shield was large enough to completely hide *E* when the stimuli were presented. *S* was seated in a straight-backed chair that he could grasp the knob with his right hand without bending his elbow. He was so positioned that his right shoulder was directly in front of the knob. In this position, the arm was not displaced to the left or right while he manipulated the knob.

S was informed that this was an experiment in torque discrimination and that his task was to turn the knob in front of him. He was told that each trial would consist of two knob turns and that he was to judge if it took more, less, or an equal amount of force to turn the knob the second time than the first. *S* was cautioned to use the equal judgment very sparingly. *E* then demonstrated and let *S* practice the correct grasp and procedure for turning the knob 90° to the right. When *S* had the knob turned the proper distance he released the knob and *E* returned it to the start position. *S* was then given 10 practice trials, 2 for each standard-variable combination. An experimental session consisted of 10 blocks of 15 trials with a 30-second rest between blocks. *S* was asked not to discuss the experiment with anyone.

When *S* used more than 5 equal judgments in any block of 15 trials, he was cautioned during the next rest-period to use the equal judgment sparingly.

After *S* released the knob, *E* stopped the rod before it could return to the start position. In this way, *S* received no cue as to the size of the stimulus from the speed or rebound of the knob.

The interval between stimuli within a trial was approximately 3 sec. The length of the intertrial interval depended upon the amount of time *S* took to make his judgment. The next trial was started about 2 sec. after *S*'s judgment.

When a session was completed, *E* totaled the number of more, less, and equal judgments. If there was a reversal between adjacent stimuli of more than 2 judgments, that condition was repeated in the next session. Reversals refer to cases where a smaller number of judgments of a given category is given to a larger stimulus. Six of the 12 *Ss* had reversals requiring repetition of at least one of the conditions.

Prior to the development of the technique finally adopted, 2 knobs were employed so that *S* had to turn one knob, then reach over 6 in. to turn the other. This procedure would have allowed more observations per unit of time by using the technique described by Shaad and Helson.[6] Unfortunately, the two-knob procedure led to an extremely large space-error that completely masked any differences in conditions; therefore, the single-knob procedure was adopted.

(5) Data. Before the data can be interpreted, the actual stimulus values must be determined. The stimuli cannot be defined in terms of loading alone for the mechanical advantage in the apparatus cannot be ignored. Stimulation must be defined in terms of the components of knob-turning. When *S* turns the knob, the force he exerts depends upon three variables—the mass of the load, the radius of the rod, and the radius of the knob. Changing any one of these values changes the amount of force *S* must exert to turn the knob.

If the apparatus is considered to be a lever, stimulus-values more appropriate than mass can be defined. The radius of the rod, k, and the radius of the knob, r, may be regarded as the two arms of a lever. Then when a load of mass m is suspended from the rod, a moment of inertia, mk, is produced. To bring the system into balance, *S* must apply to the knob another force, F, so that the resulting moment, Fr, equals mk, or:

$$Fr = mk. \qquad [1]$$

The amount of force is determined quite simply by dividing Equation [1] by r:

$$F = mk/r. \qquad [2]$$

Dimensional analysis of Equation [2] shows that the appropriate units for F are grams. *S* must apply to the knob a force equivalent to a weight of F gm. in order to bring the lever into balance. The stimuli in this study are thereby defined as equivalent to weights in grams equal to mk/r. Thus, if *S* lifts a 200-gm. load by turning the 3.81-cm. knob, the stimulus-value, F, is 33.34 gm. (200 gm. × 0.635 cm./3.81 cm.); if the load is 50 gm. and the knob has a radius of 1.27 cm., then F has a value of 25.00 gm. (50 gm. × 0.635 cm./1.27 cm.).

Because a load and its equivalent stimulus are both in units of grams, 2 different specifications will be used in presenting results. The mass of the load will always be specified in terms of grams, while the value mk/r, indicating the force, F, required to bring the lever system into equilibrium, will be specified in terms of equivalent-*gm.* (*e-gm.*). Table II lists the values of the standard stimuli in terms of *e-gm.* This table also presents the results of suspending weights from the knob to balance each of the stimuli under actual conditions of operation thus checking the theoretical against the actual values of the stimuli. The obtained values for stimuli are rounded to the nearest whole *e-gm.* There is a close correspondence between the calculated and actual values. Since the measured values are higher than the calculated values in Table II, it is evident there was a small amount of friction in the apparatus. In discussing results, the theoretical values will be used since the design of the experiment was based on them.

---

[6] Dorothy Shaad and Harry Helson, Group presentation in the method of constant stimuli as a time-saving device, this JOURNAL, 43, 1931, 422-433.

The experimental results rest upon 16,200 judgments (12 *Ss* × 9 conditions × 150 trials). Limens for greater (L_g) and less (L_l) were calculated for each of the *Ss* for each condition. The limen is defined as the stimulus-value that elicits a judgment greater or less than the *Std.* 50% of the time. The value of the liminal stimulus was determined by fitting ogives to each set of data. Separate ogives were required for the greater and less judgments. The ogives were fitted to the data using Urban's method.

TABLE II

Calculated and Measured Values of Standard Stimuli in Equivalent-grams (*e-gm.*)

Standard loading (gm.)		Knob-radius (cm.)		
		1.27	2.54	3.81
50				
	Calculated	25.00	12.50	8.35
	Observed	25	13	9
100				
	Calculated	50.00	25.00	16.67
	Observed	51	26	18
200				
	Calculated	100.00	50.00	33.34
	Observed	101	51	34

The actual computations were programmed for an IBM 1410-1401 computer and the correctness of the programming was checked by manual calculation of a limen from a representative set of values.

The point of subjective equality (*PSE*) had to be calculated before the difference limens could be determined. The method of calculating *PSE* was that proposed by Guilford according to the following Equation:[7]

$$PSE = L_1 + [s_1\ (L_g - L_1)]/(s_l + s_g) \qquad [3]$$

where L_l and L_g refer to the greater and less limens and s_l and s_g are the standard sample deviations of the two fitted ogives.

This criterion for *PSE* is also the point at which the fitted ogive for the lesser judgments crosses the fitted curve for the greater judgments. When *PSE* is interpreted in this manner, it is seen to be an indifference point separating stimuli judged greater than the standard from stimuli judged less than the standard.

Time-order-effects (*TOE*) and difference limens (*DLs*) were also calculated. The absolute *DL* can be calculated either as the difference between the limen and the standard or as the difference between the *PSE* and the limen. The latter method of computing *DLs* is very similar to that suggested by Guilford who defines the *DL* as one-half of the interval of uncertainty (*IU*).[8] The procedure used gives the same result as one-half the *IU* when the sample standard deviations are equal.

---

[7] J. P. Guilford, *Psychometric Methods*, 2nd ed., 1954, 138.

There are several advantages in using the *PSE* instead of the standard stimulus as the point from which the absolute *DLs* are calculated. The *PSE* is the stimulus-value judged to be equal to the standard and it therefore serves as a better dividing point between judgments of greater and less than does the standard stimulus. First, it enables us to measure the *DL*, a psychological variable in terms of two other psychological variables, the *limen* and the *PSE*. When the standard stimulus is used, then the absolute *DL* is defined in terms of a physical and a psychological variable. A second advantage in using *PSE* instead of the standard stimulus is, unless the sample standard deviations of the two fitted ogives are very different, the upper and lower *DLs* will be approximately the same size. When large *TOEs* occur, the method of calculating *DL* from the standard stimulus leads to asymmetry—one of the *DLs* will be considerably larger than the other. This problem does not occur when the *PSE* is used instead of the standard stimulus. Finally, a third advantage of taking *DLs* from *PSE* is that *PSE* can be related to other psychophysical data in a quantitatively meaningful way for *PSE* is a measure of the prevailing adaptation-level.

Weber fractions, or relative *DLs*, were calculated using absolute *DL* divided by *PSE* in place of the *Std.* in the denominator. The Weber fraction was taken as the unweighted algebraic mean of the greater and lesser Weber fractions.

Results and discussion. There are two different ways in which the results of this investigation can be presented. Both are equally valid and useful.

The first procedure calls for the presentation of the results in terms of grams with parameters of knob radius and loading. This technique answers the question, "When *S* turns a knob of given size and loading, what is his sensitivity to load changes?" This answer provides useful industrial engineering information. Unfortunately, when the results are presented in terms of grams, there is no way to order the nine conditions along a stimulus-dimension. This limitation makes it impossible to compare directly the results of this study with sensitivity studies under other tactile-kinesthetic conditions.

The second way of presenting the results is in terms of equivalent-grams and provides an answer to the question. "How sensitive is *S* to changes in the level of stimulation in a knob-turning task?" Answering this question provides information on sensitivity of a more general nature because the nine conditions now can be ordered on a unidimensional stimulus-continuum and are directly comparable to sensitivity studies in other modalities when translated into Weber fractions.

---

[8] Guilford, *op. cit.*, 137.

When the results are presented in terms of gram-units, then changing the value of either knob radius or loading produces concomitant changes in the size of the absolute *DL*s and Weber fractions. The effect of increasing either independent variable depends, however, upon the dependent variable being measured. Qualitatively, the results of this study may be generalized as follows: as knob radius is increased, absolute *DL* decreases and relative *DL* increases; but as load increases, absolute *DL* increases and relative *DL* decreases.

As is readily apparent, the independent variables are both directly and inversely related to sensitivity measures depending on which measure of sensitivity is employed; however, if the stimulus is specified in terms of *e-gm.*, the results can be generalized more simply: as the number of *e-gm.* required for turning the knob increases, the absolute *DL* increases and the Weber fraction decreases within the limits of this study.

TABLE III

ABSOLUTE *DL*s IN *e-gm.* AS A FUNCTION OF KNOB-RADIUS AND LOAD FOR MEN AND WOMEN

Standard load (gm.)		Knob-radius in cm.					
		1.27		2.54		3.81	
		DL_l	DL_g	DL_l	DL_g	DL_l	DL_g
50	Women	2.00	1.70	1.28	1.22	1.10	1.02
	Men	1.79	1.70	1.57	1.28	0.95	1.15
100	Women	3.51	3.42	1.93	1.92	1.46	1.35
	Men	3.35	4.65	2.25	2.24	1.54	1.59
200	Women	5.18	5.32	3.16	3.12	2.29	2.34
	Men	3.40	3.79	3.04	2.89	1.71	2.55

Let us consider the results of this experiment in detail by first examining the effect of changes in knob radius and load upon relative and absolute *DL*s.

The average absolute *DL*s are given in Table III. Without exception, when load increases, the absolute *DL* increases, as seen by reading down the columns of Table III. When this Table is read across any row, absolute *DL* is seen to be inversely related to knob radius. Whenever larger knobs are used, the size of the absolute *DL* decreases.

Statistical analysis confirms this interpretation of the results. A four-way *AOV* (sex × knob radius × loading × *DL* type) on the data of Table III shows that only knob radius ($F = 45.115$; 2/20 *df*, $p < 0.001$) and load ($F = 37.401$, 2/20 *df*, $p < 0.001$) had any effect on the size of the absolute *DL*. Sex and *DL* type were not significant at the 0.05 level nor were there any significant interactions. Trend tests were performed on the two significant sources of variance. The linear component of both knob-radius ($F = 85.479$, 1/20 *df.*, $p < 0.001$) and load ($F = 70.683$, 1/20 *df*, $p < 0.001$) accounted for the largest portion of the sums of squares in both cases. The residual component for loading was nonsignificant but for knob-radius ($F = 4.752$, 1/20 *df*) it reached the 0.05 level of significance.

Nonsignificant differences in the absolute *DLs* between the sexes may be surprising. This result may be due to the fact that the men and women in this study did not have significantly different amounts of practice in turning knobs.

The lack of differences between DL_1 and DL_g (*DL*-type) is not surprising since using the *PSE* in their calculation almost assured equality. This finding indicates that the sample standard deviations of the greater and lesser absolute *DLs* were not significantly different. If the sample standard deviations had been different, then the two absolute *DLs* would also have been significantly different. The method of calculation may, however, have been the reason for the extremely small *F* ratio (0.037) in the *AOV*. This ratio and all other *F* statistics of less than one were tested against the lower tail of the *F* distribution using the procedure given by Bennett and Franklin.[9]

TABLE IV

RESULTS OF *LSD* PERFORMED ON ABSOLUTE *DLs* ($LSD = 1.307$, $\alpha = 0.001$)

Load (gm.)	50	50	100	50	100	200	200	100	200
Knob-radius (cm.)	3.81	2.54	3.81	1.27	2.54	3.81	2.54	1.27	1.27
Force (*e-gm.*)	8.35	12.50	16.67	25.00	25.00	33.34	50.00	50.00	100.00
Absolute *DL* (*e-gm.*)	1.06	1.34	1.49	1.80	2.08	2.47	3.05	3.73	4.42

The changes in the size of the absolute *DL* caused by manipulating knob-radius and load can be simplified if the results are presented in terms of *e-gm*. From Equation [2], $F = mk/r$, it is known that increasing the load (m) or decreasing knob-radius (r) will increase the force that *S* has to exert to turn the knob. Therefore, the results given in terms of knob-radius and load (Table III) can be reinterpreted in terms of *e-gm*. As shown in Table IV and Fig. 1 when *e-gm.* stimulation is increased by using smaller knobs or larger loads, the absolute *DL* is also increased.

---

[9] C. A. Bennett and N. L. Franklin, *Statistical Analysis in Chemistry and the Chemical Industry*, 1954, **108**.

Fisher's *LSD* was determined for the nine absolute *DL*s by calculating the appropriate error term in a completely randomized *AOV* ($S_x = 0.271$, 80 *df*). Because of the effect of a prior test on the alpha level of the *LSD*, the 0.001 level of significance was chosen. Table IV shows the results of the *LSD* test. The means that are underlined by a single bracket do not differ from one another significantly. Inspection of Table IV shows that there is a progressive increase in the size of the absolute *DL* as stimulation increases. Most importantly, Table IV shows that different stimulus-values yielding the same *e-gm.* have the same absolute *DL* regardless of knob size and load. Thus a stimulus of 25 *e-gm.* produced with a 1.27-cm. knob and a 50-gm. load or with a 2.54-cm. knob and a 100-gm. load yield absolute *DL*s of 1.80 and 2.08 *e-gm.*, respectively, which are not significantly different as shown in Table IV. Similarly, the two 50-*e-gm.* conditions do not differ significantly from one another. If the *LSD* for a 95% confidence-level is calculated instead of for the 99.9% confidence-level, more means

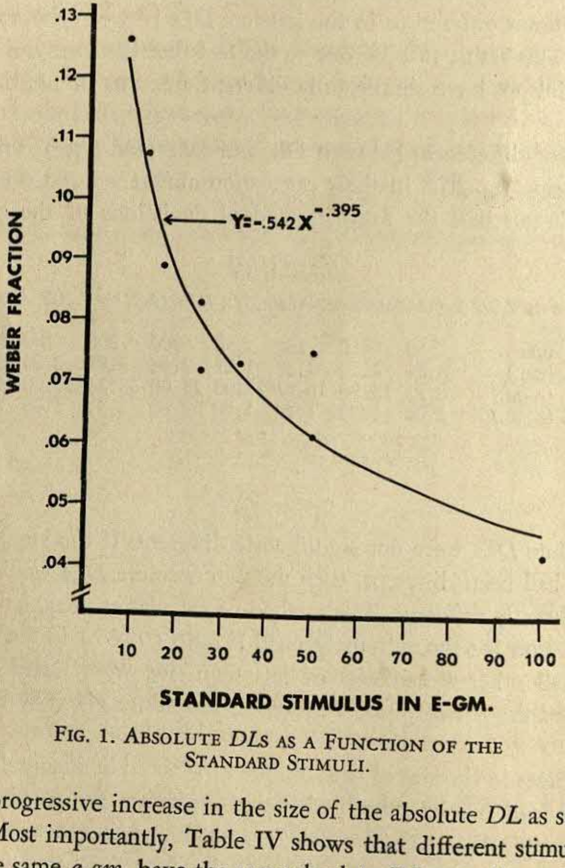


FIG. 1. ABSOLUTE *DL*S AS A FUNCTION OF THE STANDARD STIMULI.

will be significantly different and the brackets will be shorter. The duplicated stimulus-values of 25 *e-gm.* on the one hand and 50 *e-gm.* on the other still do not, however, differ significantly ($LSD = 0.761$, $\alpha = 0.05$). Thus the more important variable is stimulus-magnitude measured in *e-gm.* The significance of knob-size and load must be evaluated together; *i.e.,* in terms of actual use.

Weber fractions for each condition, with parameters of knob-radius and load, are given in Table V and Fig. 2. The relative *DL*s show, with one exception, a decrease with increased loads, as seen by reading down the columns of Table V. When read across by rows, Table V shows that the relative *DL* increases with increases in knob-radius as is to be expected since sensitivity is less with larger knobs as the hand must turn a greater amount to get the same feedback from any given load. If relative *DL* decreases with increased load, then it should increase with knob-radius, since larger knob-radius is equivalent to decreased load due to the greater lever-action of the larger knob-radius.

TABLE V

AVERAGE WEBER FRACTIONS FOR THREE KNOB-RADII AND THREE STANDARD LOADS

Standard load (gm.)	Knob-radius (cm.) 1.27	2.54	3.81	Average
	Weber fraction			
50				
Women	.074	.100	.127	.102
Men	.070	.114	.126	
100				
Women	.069	.077	.084	.082
Men	.080	.090	.094	
200				
Women	.052	.063	.069	.060
Men	.036	.059	.078	
Average	.063	.084	.096	.081

The *AOV* (sex × knob-radius × load) of the Weber fractions gives the same results as the *AOV* for the absolute *DL*s. The summary table of the *AOV* showed that only knob-size ($F = 23.837$, 2/20 *df*, $p < 0.001$) and load ($F = 13.653$, 2/20 *df*, $p < 0.001$) were significant sources of variation. Trend tests on these two sources show the linear component to be much larger than the residual component. In both cases the F values for the residual components were less than 1.00 and the linear F values were significant at $p < 0.001$.

The effect of changing knob size and load upon the relative *DL* can immediately be interpreted in terms of *e-gm.* by using Equation [2]. Weber fractions are given against stimulus-magnitude in *e-gm.* in Table VI. The underlying trend in Table VI is very apparent—as *e-gm.* values increase, the Weber fraction decreases. From Table VI it is seen that there is an almost three-fold range in Weber fractions (0.126: 0.044) as the stimulus increases twelve-fold (8.35:100).

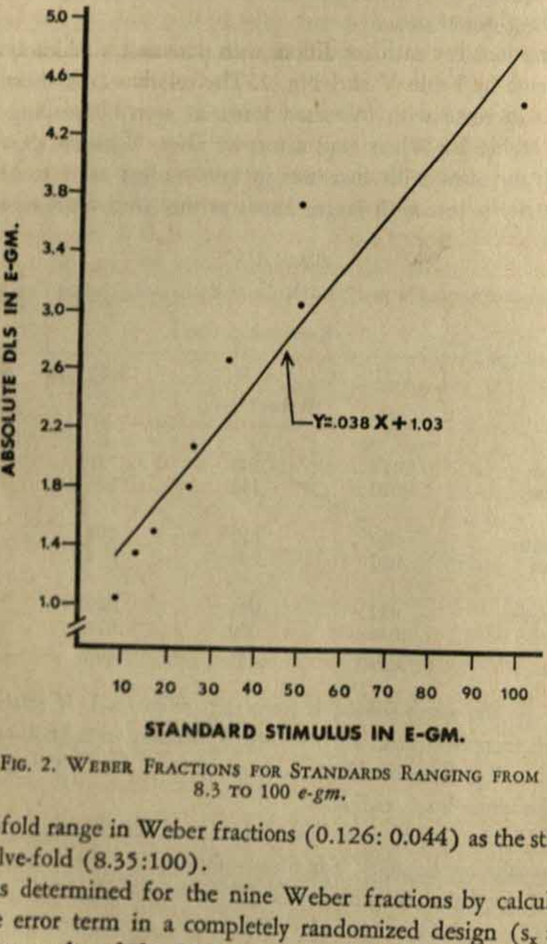


FIG. 2. WEBER FRACTIONS FOR STANDARDS RANGING FROM 8.3 TO 100 *e-gm.*

LSD was determined for the nine Weber fractions by calculating the appropriate error term in a completely randomized design ($s_x = 0.008$, 80 *df*). The results of the *LSD* are shown in Table IV. The means that are underlined by a single bracket do not differ from one another significantly.

Inspection of Table IV shows, once again, that stimulus-magnitude in *e-gm.* is the important variable because neither the two 25 nor the two 50 *e-gm.* stimuli are significantly different from each other. Nor do these conditions differ from each other significantly at an alpha level of 0.05 ($LSD = 0.023$, $\alpha = 0.05$).

Within the limits of this study, the Weber fraction is inversely related to stimulus-magnitude—a result that runs counter to the complete generality of Weber's law according to which the relative *DL* should be constant. It is well known, however, that Weber's law does not hold except for mid-range values of stimulation. Holway and Pratt reviewed sensitivity studies that had been conducted in such diverse sensory modalities as vision, audition, somesthesis, olfaction, and gustation.[10] They found that the Weber fraction decreases to asymptotic values as stimulation increases from zero and then rises sharply as stimuli values increase beyond a certain point. Woodworth and Schlosberg also report that the Weber fraction is not constant:[11] it decreases rapidly as stimulation increases from 0 to 500 gm., while between 500 and 3,000 gm. there is a comparatively small change (the region of maximal sensitivity). Relative *DLs* from stimuli beyond 3000 gm. should again increase. Since the largest stimulus in the present investigation was in the lower range of tactile-kinesthetic stimuli, it is not surprising that the relative *DL* shows a progressive and regular decrease as the stimulus-values increase from 8.35 to 100 *e-gm.*

TABLE VI

RESULTS OF *LSD* PERFORMED ON WEBER FRACTIONS ($LSD = 0.040$, $\alpha = 0.001$)

Load	(gm.)	50	50	100	100	100	200	50	200	200
Knob-radius	(cm.)	3.81	2.54	3.81	2.54	1.27	3.81	1.27	2.54	1.27
Force	(*e-gm.*)	8.35	12.50	16.67	25.00	50.00	33.34	25.00	50.00	100.00
Weber Fraction		.126	.107	.089	.083	.075	.073	.072	.061	.044

The relative *DLs* in this study show the progressive decrease reported by Holway and Pratt. The stimulus-range was not sufficient to show the typical sharp increase in relative *DL* when very large stimuli are employed. Inspection of Fig. 2 shows that maximal sensitivity is approached with stimuli as small as 100 *e-gm.* This maximal sensitivity is reached when the curve asymptotes, yielding Weber fractions of approximately 0.04. Extrapolating from the curve in Fig. 2 a stimulus of 500 or 1000 *e-gm.* will have a relative *DL* of approximately 0.04 so that it would not be inefficient to design equipment with much greater stimulus-loads than the maximum employed in this study.

---

[10] A. H. Holway and C. C. Pratt, The Weber ratio for intensive discrimination, *Psychol. Rev.*, 43, 1936, 322-340.

[11] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, rev. ed., 1954, 224.

The Weber fraction is a dimensionless measure of sensitivity that lends itself to the direct comparison of relative sensitivity in different modalities. Woodworth and Schlosberg[12] report that the value of the relative *DL* for lifted-weights of 3000 gm. is approximately 0.05.[12] Compare this fraction with the fraction obtained with the 100-*e-gm.* stimulus (0.044) and the fraction reported by Woodruff and Helson for a torque stimulus of 2146 gm.-cm. (0.044).[13] Evidently sensitivity to torque stimuli in either the lifted-rod or knob-turning task produces an asymptotic fraction of approximately 0.04 which is comparable to the asymptotic fraction for lifted-weights. There is, however, a possibility that maximal sensitivity occurs at a lower stimulus in the knob-turning task than in lifted-weight. This greater sensitivity may be the result of larger kinesthetic involvement when *S* turns a knob as opposed to lifting a weight. Certainly finger and wrist joints and muscles, used when knobs are turned, are more highly innervated and therefore capable of producing finer discriminations than are the larger arm and shoulder joints and muscles usually used in lifting weights.

The *TOE*s present a means of determining if the *PSE*s of the nine conditions were shifted differentially. A three-way *AOV* (sex × knob radius × load) yields no significant *F* ratios. A t-test, however, shows that the average *TOE* (+0.38) is significantly different from zero ($t = 2.91$, 107 *df*, $p < 0.01$).

The final area of interest was to determine if thresholds change with practice. Therefore an analysis was performed on the data collected from the two sensitivity determinations on the 2.54-cm. knob with a load series having a 100-gm. standard. These identical determinations were the first and last conditions under which data were collected. If a threshold change, caused by practice, occurs, then these two conditions should show it most distinctly.

A three-way *AOV* (sex × time × type of *DL*) was made of the greater and lesser absolute *DL*s of these two determinations. No condition or interaction yielded an *F* ratio that reached the 0.05 alpha level. Next, a two-way *AOV* (sex × time) was performed on the *TOE*s of the two determinations. The result of this analysis was the same as the first—no effect reached the 0.05 level of significance. Finally, another two-way *AOV* was performed on the Weber fractions—with the same result. Therefore, it is safe to assume that the data presented in this study do not show a systematic practice-effect. If the procedure used in measuring torque sensitivity had had a practice component, it would have shown up in the comparison between the first and last data collected under the same conditions.

---

[12] Woodworth and Schlosberg, *op. cit.*, 224.

[13] Woodruff and Helson, *op. cit.*, this JOURNAL, 73, 1965, 276.

Conclusions. When small knobs are turned against various loads the problem of determining actual stimulus-values becomes critical. Measures of sensitivity lose much of their meaning if the stimulus-magnitude is unknown or misrepresented. For this and other reasons the stimuli must be presented in terms other than that of a load of X-gm. being lifted by turning a knob of Y-cm. radius. Though under actual conditions of use, this mode of specifying stimulation can be very informative. A better method of specifying stimulus-magnitude is in terms of the force in *e-gm.S* must exert to turn the knob. This procedure allows all combinations of knob-size and loading to be ordered along one continuum. It is easy to show that the force *S* must exert to turn the knob increases with larger loads and smaller knobs.

Using this procedure of ordering stimulation, we can state, for stimuli of 100-*e-gm.* or less, that torque sensitivity is inversely related to stimulus magnitude. As stimulation increases from 8.35 to 100 *e-gm.* the Weber fraction decreases from 0.126 to 0.044.

When different knob sizes and loads produce the same stimulus value in *e-gm.*, the absolute and relative *DL*s reflect only stimulus-magnitude and not the way in which the stimulus was produced. Thus knob-radius and loading are not important determiners of sensitivity *per se*. Their importance lies in the manner in which they produce changes in stimulus magnitude as measured in *e-gm.* This conclusion does not minimize the importance of knob size because large knobs will permit loadings, much larger than investigated, to be handled easily. Furthermore, situational variables may increase the importance of both load and knob-radius.

In the region of stimulation investigated (8.35- to 100-*e-gm.*) the Weber fraction decreases rapidly in accordance with the findings of Holway and Pratt. Further increases in stimulation will not serve to further reduce the Weber fraction as it has already approached its asymptotic value.

Torque sensitivity did not show improvement with practice nor was the sensitivity of the two sexes found to be different.

DISTORTIONS IN MOVING FIGURES VIEWED THROUGH A STATIONARY SLIT

By S. M. ANSTIS and JANETTE ATKINSON, University of Bristol, England

Any figure which moves horizontally behind a stationary, vertical slit can be seen as a whole figure, although only a narrow strip of it is visible at any instant.[1] The figure appears displaced in the direction of its movement, and conspicuously compressed or foreshortened along the direction of its movement (Fig. 1A).

This effect was first noted by Zöllner who dubbed it 'anorthoscopic perception.[2] He distinguished three effects: (1) When figures are moved at a relatively high speed, they appear narrower than they really are (compressed). The figures are perceived simultaneously. (2) When the figures are moved at a relatively low speed, they appear wider than they really are (elongated). The figures are perceived successively. (3) In both situations, (1) and (2), the figures appear considerably wider than the slit behind which they move.[3]

Vierordt and Gertz studied these effects.[4] In the 1930s associates of Schumann at Frankfurt made further studies. Rothschild studied the compression during rapid pattern movements (Situation 1),[5] and Wenzel studied the elongation during slow pattern movements (Situation 2), and concluded that the two effects were of quite different origins.[6] Hecht studied the compression during rapid pattern movements, and discovered a number of new distortions.[7]

Patterns can be distorted by suitably shaped slits. Slits which are tilted, curved (Fig. 1B–F), or multiple make any moving figures appear tilted, curved, or multiple. In general, the figure will tend to conform to the form of the slit. A vertical line moving horizontally behind a slit which tilts 45° to the right will also appear to tilt to the right, increasingly as its speed of movement increases, but its tilt never quite equals that of the slit. A square moving behind a 45° tilted slit distorts into a perceived parallelogram, with its leading and trailing edges inclining toward the tilt

---

* Received for publication September 29, 1966. We gratefully acknowledge receipt of Grant No. HH/182 from the Science Research Council.

[1] T. E. Parks, Post retinal visual storage, this JOURNAL, 78, 1965, 145-147.

[2] F. Zöllner, Über eine neue Art anorthoskopischer Zerrbilder, *Ann. Phys. Pogg. Ann.*, 117, 1862, 477-484.

[3] H. L. Ansbacher, Distortion in the perception of real movement, *J. exp. Psychol.*, 34, 1944, 1-23.

[4] K. Vierordt, Der Zeitsinn nach Versuchen: Scheinbare Verzerrung bewegter Gegenstände, 1868, 191 ff.; H. S. Gertz, Untersuchungen über Zöllners anorthoskopische Täuschung, *Skandinav. Archiv. f. Physiol.*, 10, 1900, 53-73.

[5] Heinrich Rothschild, Untersuchungen über die sogenannten anorthoskopischen Zerrbilder, *Z. Psychol.*, 90, 1922, 137-166.

[6] E. L. Wenzel, Die sukzessive Erfassung der Figuren, *Z. Psychol.*, 100, 1926, 289-324.

[7] Hans Hecht, Die simultane Erfassung der Figuren, *Z. Psychol.*, 94, 1924, 153-194.

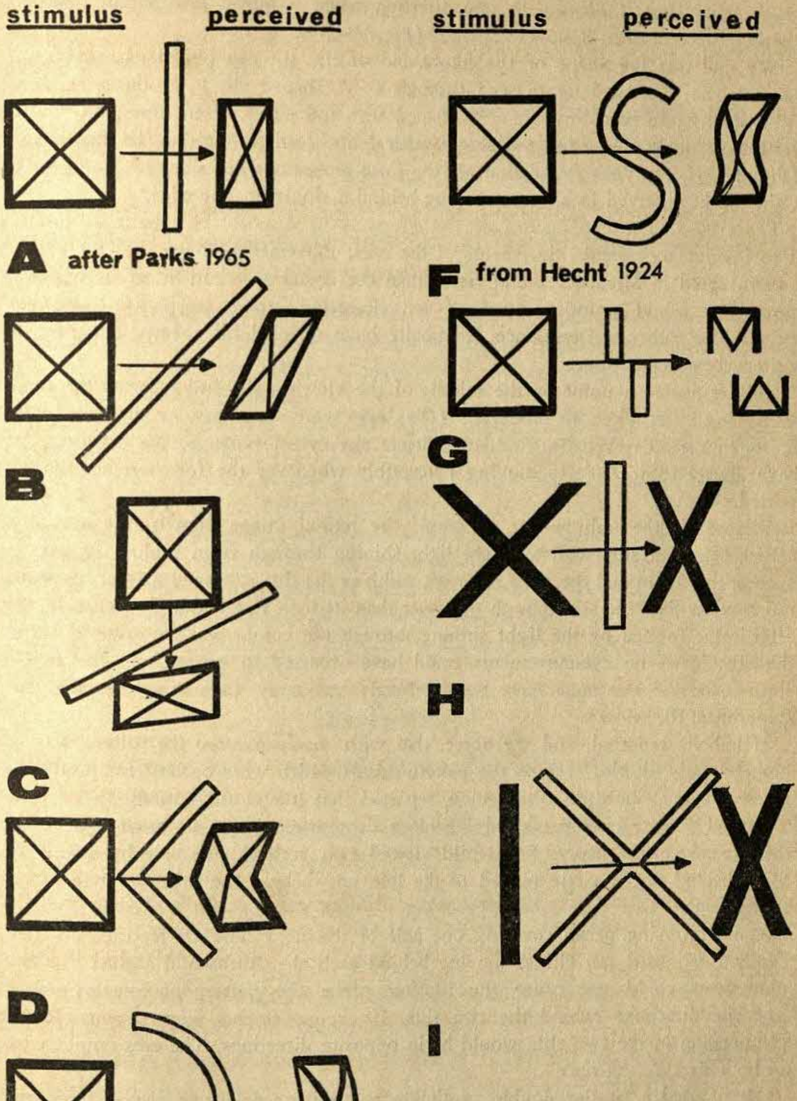


FIG. 1. FIGURES MOVING BEHIND STATIONARY SLIT
A: apparently compressed along their direction of movement.
B: conforming to the shape of the slit.

of the slit. Thus it distorts in one direction when it moves horizontally, and in another direction when it moves vertically (Fig. 1B–C).

We find that the shape of the figure and of the slit can be interchanged. For example, an 'S'-shaped figure seen through a 'V' shaped slit looks the same as a 'V'-shaped figure seen through a 'S'-shaped slit: and an 'X' seen through a single vertical slit looks the same as a single vertical line seen through an 'X'-shaped slit (Fig. 1A, H, I). Volk found that all the form alterations described by Hecht with a slit can be observed in a figure moving behind a single opaque edge.[8]

Explanation. The explanation of the effects is in dispute. Helmholtz attributed it to involuntary pursuit movements of the eyes: movements slower than the actual pattern speed in Situation 1 and faster than the actual speed in Situation 2.[9] Such movements would 'paint' compressed and elongated images, respectively, on the retina. Most other writers since Helmholtz have rejected his theory. The present writers, however, support it.

Zöllner fixated a point by the middle of the viewing slit and reported the foreshortening effect when an observer's (*O*'s) eyes were stationary, *i.e.* at times when *E*, looking at *O*'s eyeballs, could not detect any eye-movements. We disagree; we have always seen *O*'s eyes moving perceptibly whenever the foreshortening is reported.

Vierordt is the only writer to record the retinal image directly. He moved a cutout behind a slit, with a bright light shining through from behind. At certain speeds, the cutout had the same apparent width as the slit: at slower speeds, it seemed much wider than the slit, though narrower than its own real width. In each case, the after-image formed by the light shining through the cutout was the same width as the slit: hence no eye-movements could have occurred to 'spread' out the retinal image, and the eye must have been perfectly stationary (again we disagree: see Experiment IB, below).

Helmholtz reported, and we agree, that with *steady fixation* the pattern was no longer clearly visible, even at the pattern-speeds which were optimal for producing foreshortening. Vierordt, disagreeing, reported that it was still visible.

Several writers have produced ingenious demonstrations in an attempt to refute the eye-movement theory. Rothschild viewed two vertical slits side by side (Fig. 5C); behind one a circle moved to the left, and behind the other a circle moved to the right. This was done by viewing a single slit through a telescope, with a laterally reversing prism covering one half of the field. The *O*s reported an identically compressed tall ellipse moving behind each slit. Rothschild argued that eye-movements could not cause the illusion, since the pattern-movements was in opposite directions behind the two slits. If eye-movements were responsible, the compression in the two slits would be in opposite directions. The eyes cannot move in both directions at once.

Hecht used a similar double pendulum and mirror device to display two semicircles, oscillating left and right in antiphase behind a single vertical slit. The *O*s reported two tall half ellipses moving in opposite directions (Fig. 5A). He argued that no eye-movements could account for this.

---

[8] J. Volk, Tachistoskopische Untersuchungen, *Z. Psychol.*, 102, 1927, 57-106.

[9] Herman von Helmholtz, *Handbuch der physiologischen Optik*, Transl. J. C. Southall, 3, 251, 1962.

Parks also dismisses the eye-movement theory on the same grounds that two patterns can be moved in opposite directions behind the single slit without destroying the foreshortening effect.

The studies reviewed above show that the subjective compression and distortion are a function of the speed and amplitude of the figures' movement, and of the shape of the slit. The experiments reported below confirm these findings, and also show the important role played by eye-movements. They support the view that all the perceived distortion are simply due to distortions of the retinal image.

Procedure. The display was a circle or an ellipse, oscillating horizontally behind a vertical or tilted slit. *O*, while viewing this display, tracked a fixation-spot which moved in phase with the ellipse, though with a different amplitude. The fixation-spot was not occluded by the slit.

The circle or ellipse was a Lissajou figure, drawn on the tube-face of a short-persistence, Solartron CD 1014/3 oscilloscope, by a 300 c/s sine wave from an Advance signal generator. Either *E* or *O* could adjust its dimensions continuously, when necessary. It was oscillated horizontally by a slow sine wave, usually one c/s, from a Dawe VLF generator. It moved behind a narrow (2 mm × 50 mm) slit cut in a black paper mask. The fixation-spot, also driven by the Dawe VLF generator in phase with the ellipse, was on a second oscilloscope tube face, visually superimposed by a sheet of plexiglass at 45° to the line of sight.

O viewed this display binocularly from a distance of 57.3 cm. At this distance, 1 cm. subtends 1° at the eyes.

EXPERIMENT I

Eye-movements were measured during inspection of a moving figure, using after-image techniques.

(*a*) *O* viewed a circle moving sinusoidally from side to side behind a vertical slit, at frequencies ranging from 0.7 to 1.5 c/s. He reported the moment at which he first perceived a moving, tall ellipse, which occurred during the first 10-sec. observation. *E*, watching *O*'s eyes, noted that this moment coincided with the onset of *O*'s eye-tracking movements. *O* confirmed this by noting the apparent movements of a small after-image (previously printed on the fovea by means of a flash-gun): he noted that the after-image (hence, his eyes) moved with about the same frequency and phase as the circle, but with reduced amplitude: typically the after-image scanned through the 48′ visual angle, *i.e.* 8 mm., or four times the slit-width.

Incidentally, tracking movements are generally thought to require a moving object to track, yet in this situation tracking movements of 8 mm. occurred when no objects moved horizontally through more than 2 mm. (the width of the slit). This unusual result suggests that perhaps current views on the adequate stimulus for eye-tracking movements require revision.

(*b*) An experiment by Vierodt (1868) was repeated. *S* viewed a cutout (diamond, triangle, or square) in a black paper mask, which moved horizontally behind a vertical slit, with a bright light shining through from behind. Each cutout was perceived as a bright shape, narrower than its own real width but much wider than the slit. *O* then examined the after-image from each cutout. It was always a blurred replica of the perceived (compressed) shape, and was very much wider than the slit (pace Vierordt). This indicates that eye-movements were painting images on the retina, which were responsible both for the perceived shapes and for the after images.

It is concluded that eye-movements are necessary to see the distortions under discussion.

EXPERIMENT II

The subjective compression of a moving figure was measured as a function of figure- and eye-movements.

Procedure. An ellipse oscillated horizontally at 1 c/s through a distance equal to 1.3 times its own horizontal diameter. On each trial, *E* preset (1) the diameter and the amplitude of movement of the ellipse, to one of five values, as shown in Table I. This was done by adjusting the *X* gain on the oscilloscope. (2) *O*'s eye-movements to one of 3 values (12, 24, or 48 mm.). This was done by setting the fixation-spot to oscillate in phase with the ellipse, but with independent amplitude (12, 24, or 48 mm.). All 15 (3 × 5) combinations of the amplitude of the eye- and figure-movements were used.

TABLE I

DIMENSIONS OF DISPLAY IN EXPERIMENT II

Diameter (mm.)	15	20	40	100	150
Amplitude of movement (mm.)	19.5	26	52	130	195
Amplitude of movement in diameters	1.3	1.3	1.3	1.3	1.3

Since the ellipse always moved through 1.3 times its own diameter (Table I), the stimuli visible through the slit as a function of time were kept the same on all trials. Consequently, the images spread across *O*'s retinae were entirely determined by his eye-movements. The purpose of the experiment was to see whether his percepts were, in turn, entirely determined by his retinal images. *O*'s task was to indicate the apparent (subjectively compressed) diameter of the ellipse by setting its vertical diameter to apparent equality, so that the ellipse looked to him like a circle. He did this by adjusting the *Y*-gain on the oscilloscope.

Results. Results are plotted in Fig. 2 and show that the ellipse's subjective width (*y*-axis) was almost entirely dependent on *O*'s eye-movements (*x*-axis), and was almost entirely independent of its objective width and objective movement (parameters). Thus *O*'s percepts simply reflected his retinal images. The images painted on retinae had the same amplitude as his eye-movements, so we predict that if the eyes were tracking perfectly, the data points in Fig. 2 would lie on the 45° line passing through the origin, at $y = 12$, 24, and 48 mm. (equal to eye-movements). The observed widths were set at $y = 13.8$, 19.25, and 37.0 mm. These discrepancies were probably due to imperfect tracking.

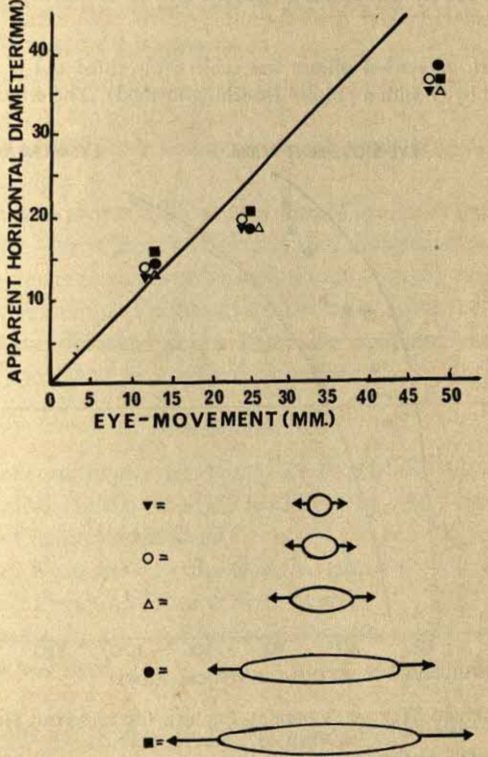


Fig. 2. Perceived Horizontal Diameter of Ellipse Oscillating Through 1.3 Times Its Horizontal Diameter

This increased with *O*'s eye-movements (y-axis) but not with horizontal 'stretching' of display (parameters). Data points are *O*'s setting. Unit-slope line shows calculated images spread across *O*'s retinae by eye-movements.

It is concluded that the subjective compression here can be attributed to the shape of the retinal image, and does not require the postulation of any post-retinal processes.

EXPERIMENTS III AND IV

Subjective tilt-distortion was measured as a function of slit-tilt, figure-movements, and eye-movements.

When a vertical figure is moved behind a stationary tilted slit, a tilt away from the vertical is induced in it. This subjective tilt was measured in Experiment III by a matching method and in Experiment IV by a null method.

Experiment III. A vertical ellipse was subjectively tilted out of vertical and the tilt was matched by *O* with a pointer (matching method). The ellipse, 4 cm. high × 1 cm. wide, oscillated horizontally at 1 c/s behind a slit which was tilted 45° to the left of vertical. Consequently the ellipse appeared to *O* to be tilted to the left. On each trial, *E* preset: (1) the amplitude of the ellipse's movement, to one of a range of 17 values lying between 22.5 mm. and 112.5 mm; and (2) the amplitude of *O*'s eye-movements. The amplitude of the fixation-spot's movement, in phase with the ellipse, was set to one of two values (24 mm. and 48 mm.).

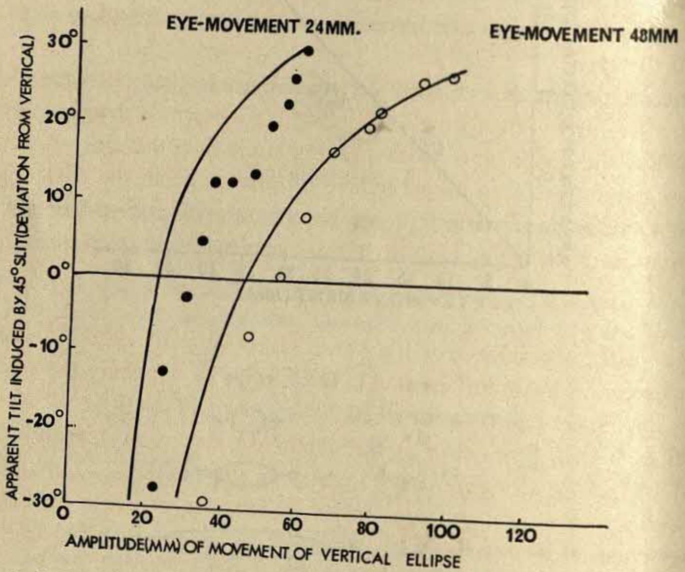


FIG. 3. PERCEIVED TILT OF VERTICAL ELLIPSE OSCILLATING HORIZONTALLY BEHIND 45° TILTED SLIT

Data points are *O*'s settings during eye-movements of 24 mm. (filled circles) and 48 mm. (open circles). Curves are calculated images spread across *O*'s retinae by eye-movements.

A plexiglass disk, free to rotate and with a pointer displayed on it, was mounted immediately in front of the slit. *O* rotated the pointer away from the vertical until it was in line with the subjective tilt of the ellipse, and its deviation from the vertical was measured by *E*.

Results. Results are plotted in Fig. 3. They show the angle of subjective tilt of the ellipse (matched by O) as a function of its amplitude of movement (set by E). The parameters of the two curves were the amplitude of eye-movements: 24 mm. (upper curve) and 48 mm. (lower curve).

The points give the observed data. The curves give the predicted data, *i.e.* the objective tilt of the images painted on O's retinae by the combinations of pattern- and tye-movements. It can be shown that the retinal tilt induced by a 45° tilted slit is given by:

$$\Phi = \tan - 1\ (R\text{-}1)/R,$$

where R is the *ratio* of the amplitude of the movement of the figures to that of the eyes.

When the figure moves with greater amplitude than the eyes—$R > 1$ and Φ is positive—the ellipse's subjective tilt is towards the slit's objective tilt. When the figure moves with less amplitude than the eyes—$R < 1$ and Φ is negative—the ellipse's subjective tilt is away from the slit's objective tilt. Insofar as the observed points lie on the predicted curves in Fig. 3, O is reporting the tilt of his own retinal images.

EXPERIMENT IV

O adjusted the subjective tilt of an objectively 40°-tilted ellipse until it appeared vertical (null method). He did this by adjusting the amplitude of its movement behind a tilted slit.

On each trial, E preset: (1) the objective tilt of the slit, within a range from 15° to 55° deviation from vertical, in steps of 5°; (2) amplitude of O's eye-movements, in the same way as in Experiment III. A random order of presentation was used.

Results. Results are plotted in Fig. 4. The points in Fig. 4. give the observed data, and the curves give the predicted data, *i.e.* the combinations of slit tilt, ellipse-movements and eye-movements that would paint a geometrically vertical image on O's retinae. Insofar as the observed points lie on the predicted curves, O perceives a subjectively vertical ellipse when his retinal image is vertical. The fit is fairly good.

The results of Experiment III and IV are similar. Taken together, they indicate that the angle of subjective tilt increases with (i) angle of slit tilt, (ii) amplitude of figure movement, and (iii) *decreasing* amplitude of eye-movements. These results are consistent with the hypothesis that O is simply perceiving the images painted on his retinae by the slit viewing conditions.

EXPERIMENT V

Subjective pattern reversals were measured when pattern- and eye-movements were opposed.

In the experiments described so far, the eyes always moved in the *same* *direction* as the figure although the *amplitude* was not always the same. In Experiment V, figure- and eye-movements were in different *directions*.

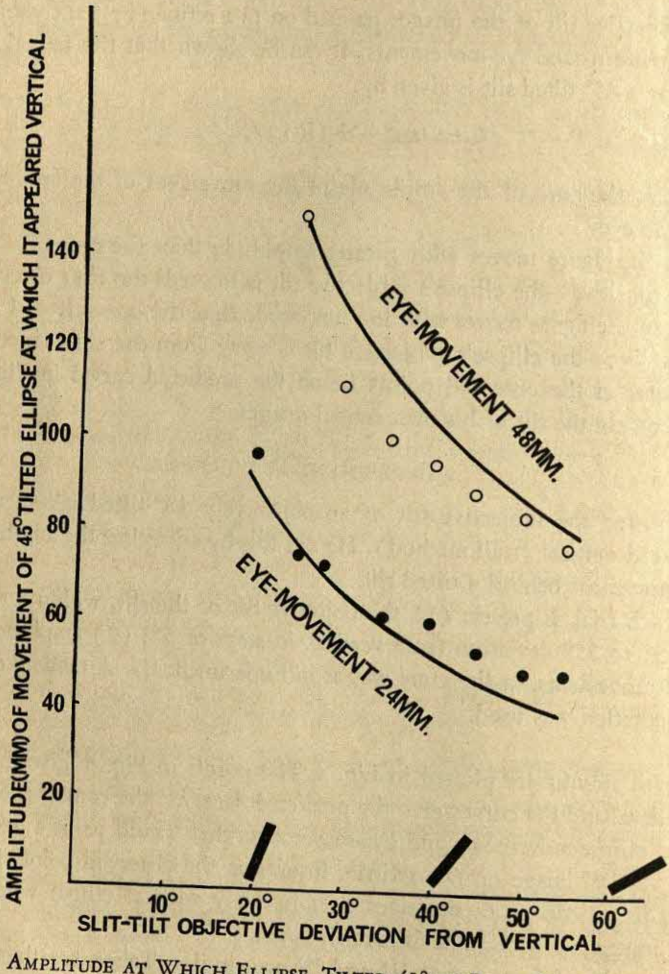


FIG. 4. AMPLITUDE AT WHICH ELLIPSE, TILTED 45° TO RIGHT OSCILLATED BEHIND SLIT, TILTED TO LEFT, TO APPEAR SUBJECTIVELY VERTICAL

Data points are *O*'s settings during eye-movements of 24 mm. (filled circles) and 48 mm. (open circles). Curves are calculated images spread across *O*'s retinae by eye movements.

Experiment III. A random order of presentation was used.

Procedure. (a) O's eyes were made to move in opposition to the figures (in antiphase) by making the fixation-spot move to the left as the figure moved to the right, and vice versa. Result: The figure then appeared reversed from left to right, as if seen in a mirror (Fig. 5, c).

(b) O's eyes were made to move at right angles to the pattern, with the slit at 45° to both. Result: The pattern was perceived as reflected through 90°, as if it were reflected in a mirror standing in the plane of the slit (Fig. 5, d).

(c) A slit was put close to a real mirror and parallel to it, and a figure moved behind the slit at right angles to it (Fig. 5, e. f). In this way, two figures were made to move in opposite directions behind two adjacent slits. The objective stimulus to *O* was a figure moving behind the real slit, with a mirror-image figure moving behind the adjacent mirror-image slit in antiphase. No fixation-spot was provided, but an after image technique showed that *O*'s eyes moved with one figure against the other.

Results. The subjective percept took one of two forms, depending on *O*'s eye-movements. The tracked figure objectively moved in the same direction as *O*'s eyes, and was retinally compressed but not reversed. The figure that was not being tracked moved against *O*'s eyes, so it was both compressed and reversed on *O*'s retina.

(i) When *O* tracked the correct-way-round figure moving to the right, he perceived two apparently correct-way-round figures moving to the right, one behind each slit (Fig. 5 e). The image in the mirror was seen the correct-way-round because it was doubly reversed (once in the mirror and once on the retina). (ii) When *S* tracked the figure seen in the mirror moving to the left, he perceived two apparently mirror-reversed figures moving to the left, one behind each slit. The correct-way-round figure was seen as reversed because it was reversed on the retina.

The figures were always seen moving in the same direction as the eyes.

These results agree up to a point with those of Rothschild and of Hecht, who found that a pair of figures moving in opposite directions behind a slit, or a pair of adjacent slits, could be seen simultaneously as compressed (Fig. 5 a.b). In our opinion, however, they misinterpreted their results. Since they used circles, which, being symmetrical, look the same whether reversed or not, they did not notice the subjective reversals, which provide strong evidence that the retinal images are reversed by *O*'s eye-movements.

Procedure. (d) A figure moved behind a slit in a transparent screen of neutral density gelatine, that it could be seen both dimly through the screen, and brightly through the slit.

Results. (1) If the eyes moved with the figure, but with reduced amplitude, the figure was seen in two different ways simultaneously: (i) veridically, through the screen, and (ii) compressed and displaced, through the

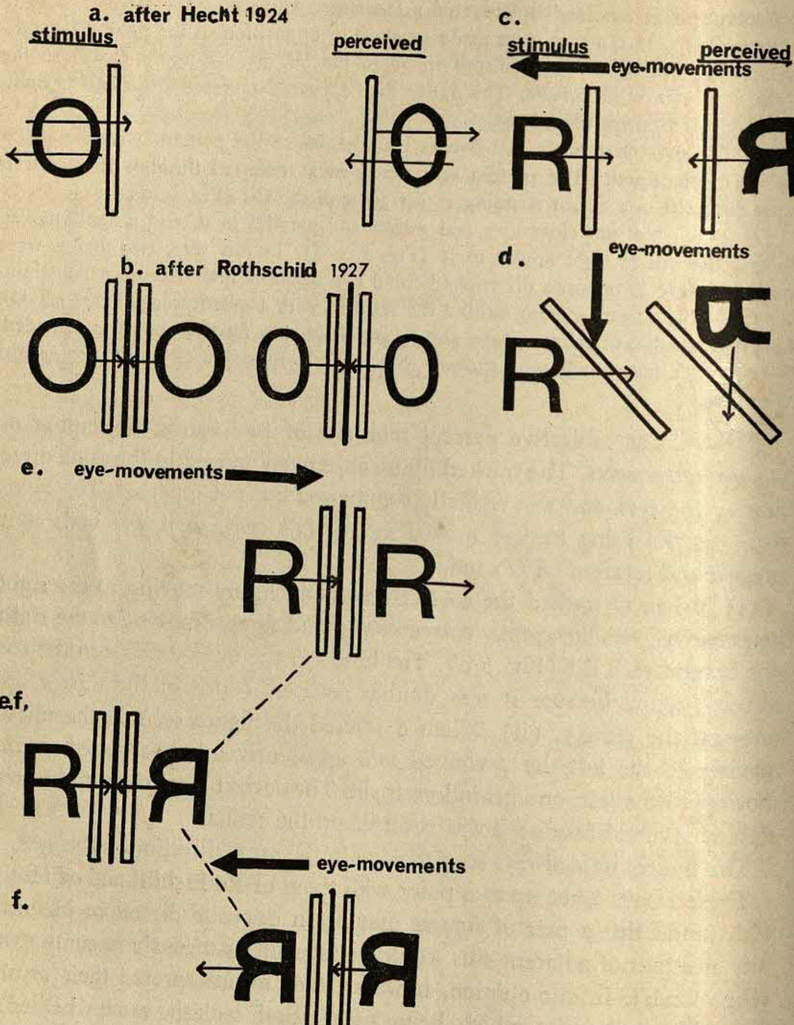


FIG. 5. PERCEIVED APPEARANCE OF TWO FIGURES MOVING IN OPPOSITE DIRECTIONS

a, b. Two figures moving in opposite directions are perceived simultaneously, apparently compressed.

c. Figure moving to the right, while eyes move to the left, is seen as mirror-reversed figure moving with the eyes to the left.

d. Figure moving to the right behind 45° tilted slit, while eyes move downwards, is seen as figure reflected through 90°, moving with the eyes downwards.

e.f. Two figures moving in opposite directions (as in b) can be seen in two ways:

e. when eyes move to the right, two correct-way-round figures are seen moving with the eyes to the right.

f. when the eyes move to the left, two reversed figures are seen moving with the eyes to the left.

slit. The effect was not unlike that of seeing a moving figure which was lit both by a dim, continuous light, and by a brief, bright flash.

(2) If the eyes moved against the figure, it was again seen in two different ways simultaneously: (i) Veridically through the screen, and (ii) compressed, reversed left to right, and moving with the eyes, through the slit. Our interpretation is that the figure was seen veridically through the screen, while the bright, compressed image was briefly painted on the retina when eye-movements made the slit scan the retina.

EXPERIMENT VI

Distortions of depth were measured in Experiment VI. When a flat, 2D picture moved behind a slit, it was shown that its retinal image is compressed when the eyes move in the same directions (Experiments I and II), and is reversed when the eyes move in the opposite direction (Experiment V).

In Experiment VI, a three-dimensional (*3D*) object was used instead of a two-dimensional picture. In practice, a *3D* object cannot conveniently be viewed through a single stationary slit, since parallax causes each eye to see a different 'slice' of the object. For this reason, an anaglyph stereogram was used, consisting of a pair of superimposed stereo pictures, one in red and the other in blue. The left eye looked through a blue filter and the right eye through a red filter, each seeing one picture. The two eyes together saw 'depth' owing to the disparity between the red and blue pictures.

The stereogram figures were cut out of 'Letrafilm' self adhesive plastic film No. 109 (pale blue) and 145 (pink) and viewed through Strand Electric 'Cinemoid' No. 19 (dark blue) and No. 6 (primary red). The stereogram was oscillated horizontally, close behind a stationary vertical slit. *O*'s eyes were made to track in phase with the object by means of a moving fixation-spot.

Results. (a). The figure appeared to be compressed laterally from left to right and also in depth. This is presumably because the retinal disparity was also laterally compressed. The subjective effects were the same as if the retinal images from the stereograms had actually been laterally compressed. (*b*). *O*'s eyes were then made to track a moving spot to the left while the stereogram moved to the right and vice versa. Under these conditions, the whole figure appeared laterally reversed as in a mirror and depth was also reversed, *e.g.* a hitherto distant object on the left, appeared as a near object on the right. (*c*) If the stereogram was oscillated vertically behind a horizontal stationary slit and *O*'s eyes tracked in the same plane but in antiphase, the object appeared mirror reversed, up-down, but depth was unchanged, *i.e.* not reversed. The vertical movement reversed the disparity on the retina up-down, but not left-right so the retinal images did not give reversed depth.

It is concluded that these results are entirely determined by the retinal images.

Summary and conclusions. The experiments described here indicate that the perceived patterns correspond to the images painted on the retina by eye-movements.

These movements spread the image of the slit, and of the moving figure behind it, across the retina. The figure is effectively scanned by the slit, owing to the figure's movement behind the slit, and reconstituted on the retina, owing to eye-movements. Any asynchrony between pattern- and eye-movements will cause geometrical distortions in the retinal image, just as any asynchrony between a TV camera's scan and a TV receiver's scan will cause distortions of the picture on the TV screen. If the eye-movements are of smaller amplitude than pattern movements, then the image formed on the retina is geometrically compressed along the direction of movement. If the eye-movements are the same as the figure-movements, corresponding to perfect tracking by the eye, then the retinal image is perfect. If the eye-movements are of greater amplitude than the figure-movements, the retinal image is geometrically elongated along the direction of movement. Furthermore, if the slit is tilted or curved, then any asyncrhony between eye- and figure-movements imparts a tilt or curvature to the retinal image. A *tilted* slit cuts the figure into tilted 'slices,' and these slices are compressed (not necessarily at right angles to the slit) along the direction of eye- and figure-movements. When eye- and figure-movements are in different directions, the compression is along a 'compromise axis.' If the eye-movements are of smaller amplitude than the figure-movements, then the retinal image is geometrically tilted toward the direction of the slit. If the eye-movements are identical to the figure-movements, corresponding to perfect tracking, the retinal image is identical to the stimulus-figure. If the eye-movements are of greater amplitude than the figure-movements, the retinal image is geometrically tilted in a direction opposite to the tilt of the slit.

All these effects are purely optical, not psychological, and can be reproduced with a camera. They are often visible in photo-finish pictures, where the film is drawn past a stationary vertical slit, and in photographs of fast cars taken by a camera with a moving-slit focal-plane shutter, such as a Leica. Indeed the effect pre-dates photography. In 1824, Peter Roget, the versatile author of the *Thesaurus,* noticed in Bristol that the wheel spokes of cabs appeared curved, convex downwards on both sides, as they moved past the slits between the vertical palings of a fence. He noticed how "the successive appearance of the different parts of a spoke, as the wheel moves along and the spoke moves up or down, would combine to give the total impression of a curved spoke, a curve for which he wrote the formula."[10] This is essentially the same effect as the one examined in this paper.

---

[10] E. G. Boring. *Sensation and Perception in the History of Psychology*, 1942, 588f.

FACTORS INVOLVED IN JUDGMENTS MADE BY THE METHOD OF CONSTANT STIMULI

By A. J. GILBERT, University of Western Australia

Helson claims that with the method of 'comparative' judgments the *PSE* is a weighted geometric mean (comparative adaptation-level or *CAL*) of the standard stimulus and the adaptation-level to the series of comparison-stimuli (series adaptation-level or *SAL*).[1] Unfortunately, however, his equation gives relatively limited insight into the process by which such judgments are made.

The procedure employed in the present experiment involved pressing two parallel bars of a standard stimulus onto the anterior surface of one of the *S*'s forearms at the same time as the bars of one of a series of similar comparison-stimuli were pressed onto the anterior surface of the other forearm. *S* was asked to say whether the bars of the comparison-stimulus were further apart or closer together than the bars of the standard. Two levels of temporal interval between successive judgments were employed, and one standard stimulus with three levels of comparison-series. Only one comparison-series was used with each *S*, who was adapted initially by application of both the standard stimulus and the relevant comparison-stimuli. Each *S* then made another set of judgments with the loci of application of the standard and the comparison-stimuli interchanged to determine whether separate adaptations were involved, one specific to the standard and the other specific to the comparison-series. The experiment was designed to show also whether the results were influenced by the particular forearms to which the standard and the comparison-stimuli were applied.

Method: (1) Apparatus. The apparatus included a standard stimulus and seven stimuli for comparison similar to those described by Day and Singer.[2] Each stimulus had two parallel Perspex blades with their 3.25 × 0.15 cm. lower surfaces in the same plane. The distance between the blades was 11.0 cm. on the standard and varied from 6.5-15.5 cm. by 1.5 cm. steps on the comparison-stimuli.

---

* Received for publication November 3, 1966. The author is indebted to Dr. J. Ross and to Dr. A. J. Marshall, both of the University of Western Australia, for advice on the preparation of the report, and to the Education Department of Western Australia for the use of testing facilities at Claremont and Graylands Teachers Colleges.

[1] Harry Helson, *Adaptation-Level Theory*, 1964, 220.

[2] R. H. Day and G. Singer, A tactile spatial aftereffect, *Aust. J. Psychol.*, 16, 1964, 33-37.

(2) Subjects. The *Ss* were 120 students from Claremont and Graylands Teachers Colleges, Western Australia. All of the *Ss* were volunteers and *right handed.* Approximately one quarter of them were men.

(3) Design. The first 72 *Ss* were assigned to Group 1 and the other 48 *Ss* to Group 2. The interval between judgments was 5 sec. for Group 1 and 10 sec. for Group 2, this being the only difference between the treatments for the two groups. Three levels of comparison-series were employed—6.5-12.5, 8.0-14.0, and 9.5-15.5 cm.—each level comprising five stimuli. The *Ss* were randomly allocated to comparison-series, 24 per level in Group 1 and 16 per level in Group 2.

(4) Procedure. Individual *Ss* were blindfolded and seated with their forearms resting supinated on a table. The standard was manually placed with its lower surfaces resting on the anterior surface of one forearm for approximately 1 sec., and one of the comparison-stimuli was simultaneously placed on the other forearm in the same manner. The pressure on the stimulated areas would have been approximately 454 gm. per sq. cm. if the contact had been even over the surfaces of the blades, but the curvature of the forearms prevented completely even contact. The blades were equidistant from the mid-points of the forearms, and *S* was asked whether the blades of the comparison-stimuli were further apart or closer together than those of the standard. Each *S* made 100 judgments at one session. The five relevant comparison-stimuli were presented in random order in 'rounds' of five presentations each, and each stimulus was presented once in each round. The standard was applied to the right arm for the first 60 judgments for half of the *Ss* in each group and to the left arm for the rest of the *Ss*. Following the 60th judgment, the standard and the comparison-stimuli were interchanged in respect to the arm of application. The interval between the 60th and the 61st judgments was 20 sec.

Results. Derived *PSE*s from each of the 12 sub-groups over Rounds 5-12 were calculated (least squares) and these were averaged over the various levels because of systematic differences between the *PSE*s from the sub-groups. In each case, however, the averaged *PSE*s were practically identical with those from pooled results. Theoretical *AL*s were calculated by Helson's equation number 1A setting $K = O$, $k_1 = 3$, and $k_2 = 1$.[3] This equation takes into account an anchor or comparison (standard) stimulus but is not the equation for the method of comparative judgments.[4]

The effects of the comparison-series over rounds are shown in Fig. 1 in terms of the responses to the common (9.5, 11.0, and 12.5 cm.) stimuli. As can be seen, there was an initial adaptation to the comparison-series and then reasonably stable responses up to the interchange between arms. The *PSE*s (Rounds 5-12) were 9.67, 10.87, and 11.96 cm., and the *AL*s were 9.66, 10.84, and 11.975 cm., respectively. The largest differences between the *PSE*s and the relevant *AL*s was 0.03-cm. Over Round 13 (post interchange) the results were close to the same as those from the first round. The results from the three series then diverged once more. Although the extent of these differences appeared to be independent of the interval between judgments, the significance was calculated for the two groups separately and in terms of the total responses to the three common comparison stimuli. The differences from Rounds 5-12 were highly significant (5 sec., $F = 17.086$, $df = 2/66$, $p < 0.001$; 10 sec., $F = 31.049$, $df = 2/42$, $p < 0.001$); those from Rounds 17-20 also were significant, but at a lower level of probability (5 sec., $F = 4.581$, $df = 2/66$, $p < 0.02$; 10 sec., $F = 7.544$, $df = 2/42$, $p < 0.01$).

---

[3] Helson, *op. cit.*, 129.

[4] Harry Helson, W. C. Michels, and Artie Sturgeon, The use of comparative rating scales for the evaluation of psychophysical data, this JOURNAL, 67, 1954, 321-326.

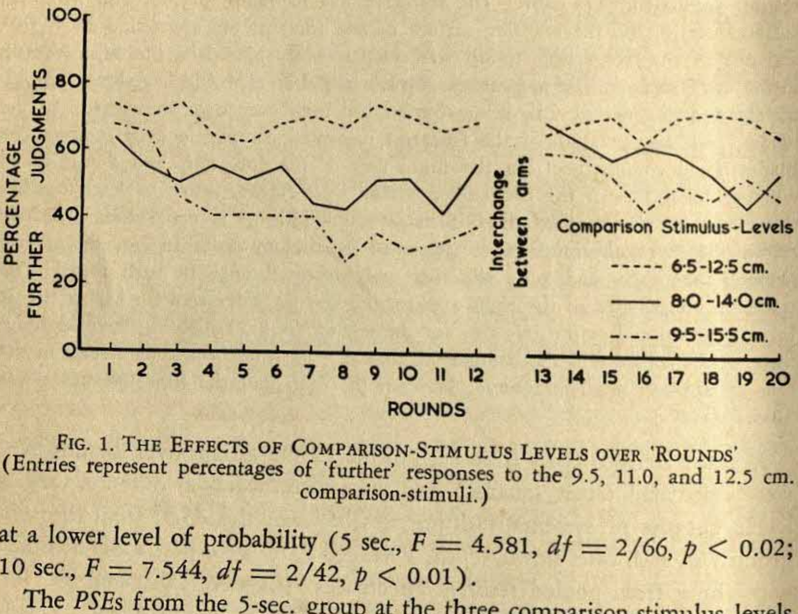


FIG. 1. THE EFFECTS OF COMPARISON-STIMULUS LEVELS OVER 'ROUNDS'
(Entries represent percentages of 'further' responses to the 9.5, 11.0, and 12.5 cm. comparison-stimuli.)

The *PSEs* from the 5-sec. group at the three comparison-stimulus levels were 10.04, 10.69, and 11.825 cm., and those from the 10-sec. group were 9.20, 11.055, and 12.095 cm., respectively. As may be seen from Fig. 2, it appears there might have been some relation among the *PSEs*, the *ALs*, and the arithmetic means ($\overline{X}$s) of the comparison-stimuli. It seems significant that the lines joining the *PSEs*, the *ALs*, and the $\overline{X}$s between the two smallest series, intersect at about the same point.

The *PSEs* from the 6.5-12.5, 8.0-14.0, and 9.5-15.5-cm. series (*i.e.* averaged over intervals) were 9.47, 10.69, and 11.465 cm. from the right-arm *Ss*, and 9.87, 11.06, and 12.46 cm. from the left-arm *Ss*, respectively. As may be seen from Fig. 3, these *PSEs* fell symmetrically around the *ALs*, and again it seems there might have been some relation among the *PSEs*, *ALs*, and $\overline{X}$s.

As illustrated in Fig. 4, there were differences depending both on the arms to which the stimuli were applied and on the interval between judgments. With the 5-sec. interval, the right-arm *S*s over Rounds 1-12 gave an average of 22.2% more 'further' responses to the three common stimuli than did the left-arm *S*s ($F = 12.812$, $df = 1/66$, $p < 0.001$). The difference with the 10-sec. interval was in the same direction but only 4.0% and insignificant ($F = 0.333$, $df = 1/42$). After the interchange between arms, the 5-sec. intervals *S*s for whom the comparison-stimuli *then* were being applied to the right arm, gave 2.8% fewer, further 'responses' than the *then* left-arm *S*s, the difference being insignificant ($F = 0.136$, $df = 1/66$). On the other hand, although there was only an unreliable difference between the results from the right- and left-arm 10-sec. *S*s over the first four rounds after the interchange (6.25% more 'further' responses from the *then* right-arm *S*s), an increasing difference appears to have developed over the last four rounds.

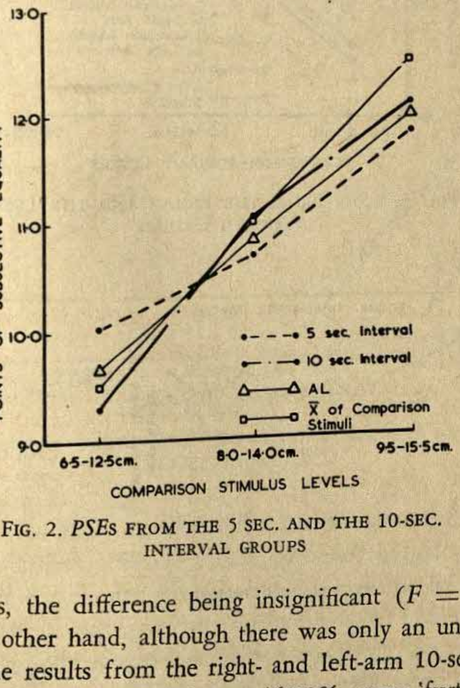


FIG. 2. *PSE*s FROM THE 5 SEC. AND THE 10-SEC. INTERVAL GROUPS

The *PSE*s from the right- and left-arm sub-groups (Rounds 5-12) are illustrated in Fig. 5. The differences between the *PSE*s were approximately the same at each comparison-level with the 5-sec. interval, while they appear to have varied over levels with the 10-sec. interval .

Discussion. The original hypothesis was that with the method of constant stimuli *S* would develop two more or less separate *AL*s, one to the

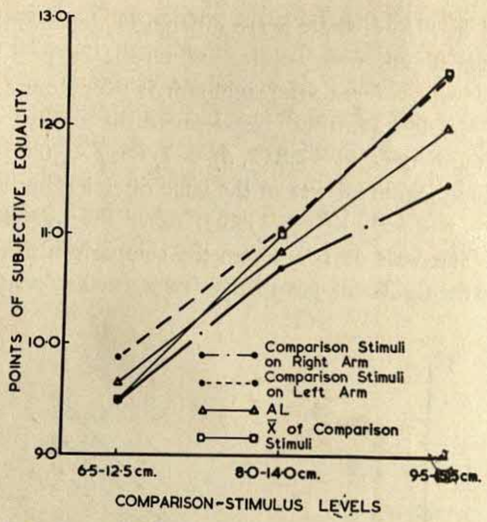


FIG. 3. *PSE*s FROM THE RIGHT- AND THE LEFT-FOREARM GROUPS

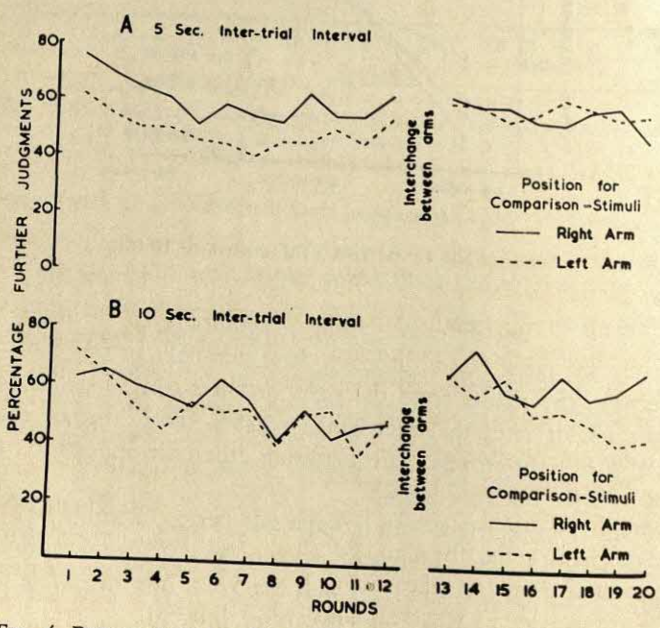


FIG. 4. PERCENTAGES OF 'FURTHER' RESPONSES TO THE 9.5, 11.0, AND 12.5 CM. COMPARISON-STIMULI FROM THE RIGHT- AND LEFT-FOREARM GROUPS SEPARATELY OVER 'ROUNDS'

standard and the other to the comparison-series, but the situation was more complex and it seems at least that a third factor must be assumed. The adaptive effects from the three series indicate there must have been at least two factors. If there had been only one factor against which both the standard and the comparison-stimuli were judged, the *PSEs* should have been 11.0 cm. On any particular occasion, each stimulus would have been compared with a single *AL* and the response would have been in terms of a comparison between these two relative values. Such responses should have been at least directionally correct, and a comparison-stimulus objectively equal to the standard should have elicited an 'equal' judgment.

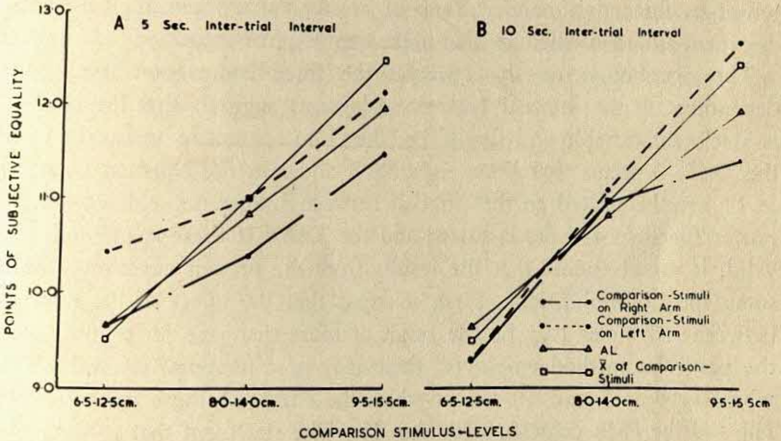


FIG. 5. *PSE*s FROM THE RIGHT- AND LEFT-FOREARM GROUPS AT THE 5-SEC. AND 10-SEC. INTERVALS

The trends of the results after the interchange between arms suggest there must have been at least a third factor involved. Had there been only two factors, the results should have shown an interaction in that with the 6.5-12.5 cm. series, the percentage of further responses to the common comparison-stimuli should have reversed from above 50% over round 12 to below 50% over Round 13, and vice versa with the 9.5-15.5-cm. series. For example, when the 11.0-m. standard was applied to an arm adapted to the 6.5-12.5 cm. series the standard should have been judged as relatively high in value. Similarly, the stimuli from the 6.5-12.5 cm. series should have been perceived as relatively small when applied to an arm adapted by application of the standard. In fact, the trends immediately following the interchange were similar to those from the initial rounds. Although it might be argued that this apparent readaptation was due to the 20-sec. break in testing, a previous experiment with the same apparatus, similar procedures, and a 2-min. rest *without* an interchange between arms, showed, if anything, slightly greater adaptation-effects immediately after the rest than before it.[5] It seems that this third factor might have been a general adaptation to all of the stimuli.

The differences between the *PSE*s depending on the arm to which the stimuli were applied suggest that some factor probably associated with cerebral dominance influenced the judgments. A stimulus tended to be perceived by these right-handed *S*s as of greater value when applied to the dominant arm than when applied to the non-dominant arm.

The distribution of the *PSE*s at the three comparison levels and depending on the interval between judgments suggests that the interval is a relevant variable and should be taken into account in an equation for the *CAL*. It seems that the weightings s and t in the equation $CAL = S^{s}A'^{t}$ may be related to the interval between judgments and not simply power functions for the standard and the *SAL*.[6] If these suggestions are valid, it would appear that the results from the present experiment have some specific implications. First, it seems that the effect of the interval between judgments may be the result of more than one factor and, with the particular method employed, there may be a temporal interval somewhere between 5 and 10 sec. at which the effects of these factors cancel out and the *PSE*s coincide with the *AL*s. The statement that $s + t = 1$ would support this.[7] Secondly, the coincidence of the intersections of the lines joining the *PSE*s, the *AL*s, and the $\overline{X}$s from the 5 and the 10 sec. groups between the two smallest comparison-series (Fig. 2) suggests that, if the $\overline{X}$ of the comparison-stimuli equals the *AL*, the *PSE* may be equal to the *AL* and independent of the interval between judgments. Thirdly, with increase or decrease in the length of the interval, the *PSE*s may approach limiting values which could yield information concerning the nature of the underlying functions.

In general, it seems that, rather than being given by a simple geometric equation, the function for the *CAL* must make reference to three adaptive factors and the interval between successive judgments. Furthermore, it appears possible that the effect of the interval between judgments is a function of more than one factor.

---

[5] A. J. Gilbert, Tactile spatial after-effect or adaptation-level, *J. exp. Psychol.*, 73, 1967, 450-455.

[6] Helson, Michels, and Sturgeon, *op. cit.*, 324.

[7] *Ibid.*, 324.

SUMMARY

It has been suggested that with the method of constant stimuli the *CAL* is a weighted geometric mean of both the standard stimulus and the *AL* to the series-stimuli. The present results for tactile stimulation of the forearms indicates that there may be several factors involved in making such judgments, and that derived *PSEs* probably approximate the *CAL* only when the stimulus-values or the intervals between judgments are such that some of the effects balance out. The factors may include general adaptation, a specific adaptation to the standard stimulus, adaptation to the comparison-stimuli, and some factor or factors related to the interval between judgments. Furthermore, it seems that with the particular method employed there was a relatively constant factor related to cerebral dominance.

THE EFFECT OF TRAINING PROCEDURES ON THE DOUBLE ALTERNATION-BEHAVIOR OF LABORATORY RATS IN A TEMPORAL MAZE

By LEONARD DIAMOND, Camarillo, California

In a temporal maze, there is only one choice point at which the animal is required to make either a right or left turn; either one bringing it back to the same choice point for another turn. The temporal maze is to be contrasted with the spatial maze where each choice-point is spatially different.

Past studies indicate that humans, monkeys, raccoons, dogs, and cats have all performed significantly in the temporal maze solving the double alternation problems.[1] Some investigators have reported demonstrations of double alternation behavior in the laboratory rat; however, these studies utilized spatial mazes, water mazes, or a modified form of the Skinner Box.[2] Whenever double alternation in the temporal maze has been demonstrated, or has failed to be demonstrated, the interpretation has been in terms of symbolic processes. Such processes are assumed to increase in complexity with ascension in the phylogenetic scale. Failure to find double alternation in the rat has led to the conclusion that this animal is perhaps the lowest point in the scale at which symbolic processes appear, or above which they do appear.

Observations of the rat learning a spatial maze of whatever kind have been interpreted by some in terms of the chaining of stimulus-response connections.[3] Others have tended to interpret the results in terms of place learning.[4] It would appear that if the rat is given the opportunity to learn a place (as in an elevated, open maze), it does so in preference to learning a chain of responses.[5] This latter

---

* Received for publication November 28, 1966. From Camarillo State Hospital. Experiment was performed at The Catholic University, Washington, D.C.

[1] H. W. Karn and R. A. Patton, The transfer of double alternation behavior acquired in a temporal maze, *J. comp. Psychol.*, 28, 1939, 550-561; W. S. Hunter, The behavior of raccoons in a double alternation temporal maze, *J. genet. Psychol.*, 35, 1928, 374-388; W. S. Hunter and S. C. Bartlett, Double alternation behavior in young children, *J. exp. Psychol.*, 38, 1948, 558-567; L. W. Gellerman, The double alternation problem, I, II, III, *J. genet. Psychol.*, 39, 1931, 50-72, 197-226, 359-392.

[2] W. S. Hunter, The sensory control of the maze habit in the white rat, *J. genet. Psychol.*, 36, 1929, 505-537; Harold Schlosberg and Arnold Katz, Double alternation lever-pressing in the white rat, this JOURNAL, 56, 1943, 274-282; W. S. Hunter and J. W. Nagge, The white rat and the double alternation temporal maze, *J. genet. Psychol.*, 39, 1931, 303-319; Sol Evans, The role of kinaesthesis in the establishment and control of the maze habit, *J. genet. Psychol.*, 48, 1936, 177-198.

[3] K. W. Spence, The nature of discrimination learning in animals, *Psychol. Rev.*, 43, 1936, 427-449; Eugene Galanter and W. A. Shaw, "Cue" vs. "reactive inhibition" in place and response learning, *J. comp. Physiol.*, 47, 1954, 395-398.

[4] E. C. Tolman, The determiners of behavior at a choice point, *Psychol. Rev.*, 45, 1938, 1-41.

[5] H. F. Gallup and Leonard Diamond, Transfer of double alternation behavior of rats in a temporal maze, this JOURNAL, 73, 1960, 256-261.

type of learning is utilized in enclosed mazes where the animal does not have the opportunity for locating the goal in space. This response chaining is presumably a series of stimulus-response connections, where the stimuli are primarily external ones from the maze, and possibly internal ones from the animal's own proprioceptive feedback. That is, kinesthetic impulses arriving in the central nervous system from one response may serve as cues for the next response in a chain.

Hunter's work dealt mainly with double alternation and delayed response. He explained double alternation of the raccoon in terms of symbolic processes and indicated that the rat lacks the capacity to establish learning on a symbolic level. Hunter also suggested that an animal can learn to double alternate through "the accumulation of internal events."[6] This is taken to be a chaining of responses based upon kinesthetic stimulation.

Schlosberg and Katz concluded that symbolic activity is "the cumulative effect of previous behavior, stimulus-traces, continuing brain-processes, and fractional responses."[7] This seems to include the concept of symbolic processes as one kind of response-chaining. Hunter defined symbolic processes as "substitute stimulus-response relationships that can be retained and later recalled to guide behavior selectively."[8] This definition seems to exclude a straight-forward chaining of responses; at least it emphasizes the substitute aspect of "representational mediation-process," as Osgood stated.[9]

An experiment by Gallup and Diamond was designed to provide a test between these two hypotheses.[10] It was demonstrated that rats can learn to double alternate under specific conditions of training and that they are able to transfer this training to other mazes which require different muscular responses thus supporting the hypothesis of symbolic processes. The significant performance on the part of the original experimental animals suggests three explanations: (1) Exploratory behavior in the second maze interfered with the demonstration of double alternation behavior; (2) the training technique used initially is superior to that used in the transfer training, which was only the fourth stage of the original training, and (3) the number of training trials was different in the two training sessions.

The present study has been designed to test the second and third questions posed by the previous experiment. By keeping all of the previous variables constant and manipulating only the training techniques and number of trials, it is hoped that the reasons behind successful double alternation may be uncovered. The hypotheses to be tested are: (a) There will be a significant difference in the double alternation behavior of rats in a temporal maze that can be attributed to a specific training procedure; and (b) the training procedure that best demonstrates double alternation behavior will uphold the theory that this type of behavior is based upon a symbolic process.

---

[6] Hunter, *op. cit.,* 1928, 374.

[7] Schlosberg and Katz, *op. cit.,* 1943, 274.

[8] Hunter, *op. cit.,* 1929, 505.

[9] C. E. Osgood, *Method and Theory in Experimental Psychology,* 1953, 658.

[10] Gallup and Diamond, *op. cit.,* 1960, 256.

Subjects. The *Ss* consisted of 20 male Canadian Hooded Rats bred at the Walter Reed Army Medical Center Animal Experimental Laboratories in Washington, D.C. The *Ss* were 84 days old at the beginning of the experiment. They were all given 10 days of preëxperimental handling by *E* without gloves. During this gentling period, they were allowed to run freely over a table top and to climb on blocks of wood and wire mesh while being handled. This procedure was carried out for 1 hr. each day. The *Ss* were then randomly divided into two groups and their tails were marked with a band of indelible ink. In addition they were marked with letters as follows: Group 1 consisted of rats, k–t; Group 2, of rats, a–j.

Apparatus. The temporal maze was so designed that the animal could not learn to go to a particular place; instead, it was forced to pass the place where it had received food-reward many times during each trial. The maze, as shown in Fig. 1, is without external differentiating stimulus-cues to guide behavior at the choice-point. Rather than guiding behavior, the maze can only serve to interfere with the animal's kinesthetic sensations. The animal in this maze cannot utilize external stimuli, hence it must make all responses to stimuli within itself, ignoring any external maze-stimuli.

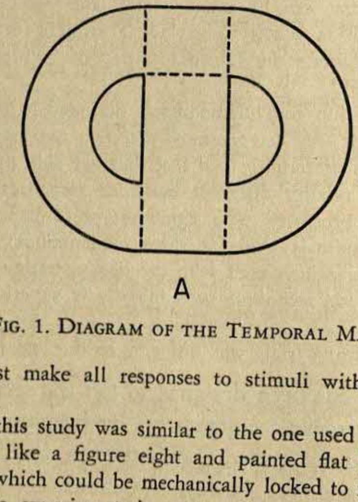


FIG. 1. DIAGRAM OF THE TEMPORAL MAZE

The maze used in this study was similar to the one used in the previous experiment. It was shaped like a figure eight and painted flat black. It was arranged with swinging doors which could be mechanically locked to force the animal during training to follow the experimenter's pattern in order to find food reward. The maze was 42 in. in length, 30 in. in width, and 15 in. deep. The actual running alleys were 4 in. wide, making it difficult for the animals to turn around during training trials. The maze was constructed of pine and sheet metal and mounted on a solid pine base which was bolted to a table. The swinging doors were constructed of light metal to facilitate the animal's pushing them open. All doors had four ¼-in. holes cut through that the animal could see into the next alley. Doors could be locked by sliding a metal rod through special brackets.

Lighting was kept constant during all trials with a 60 w. bulb suspended over the center of the maze. The maze was kept in the same room of the vivarium where the animals were housed. Lighting, temperature, and noise-level were kept constant throughout all training and test-trials. Time of trials was kept on a standard stop-watch and entered on prepared data-sheets.

Method and procedure. Each group of animals was studied. Before training in double alternation, all the animals (20) were equated in time-scores received from running a straight alley 28 in. in length. These scores were recorded, the rats were ranked, and the two groups were formed. In this way, the need for 20 additional control animals was alleviated. The paired animals were so housed that in any one cage one animal from each group could be found. These animals were matched as to deprivation schedule, number of trials, and handling.

Group 1 was trained as follows: (*R*) (*RR*) (*RRL*) (*RRLL*) or (*L*) (*LL*) (*LLR*) (*LLRR*). Group 2 was trained: (*RRLL*) or (*LLRR*). All animals in Group 1 received 20 trials on the first turn, followed by 25 trials on the second turn, 20 trials on the third turn, and 20 trials on the fourth turn. Training was thus completed in four stages, each stage adding a new turn until the *RRLL* or *LLRR* patterns had been completed. The deviation in the number of trials in the second stage of learning was necessitated by the fact that this was the hardest stage for the animals to learn. The second turn required the animals to pass the area where they had previously received food reward. Once this had been accomplished, the remaining stages were most easily learned and 20 trials for each stage were considered to be sufficient.

The animals in Group 2 were given the same number of training trials (85) in the maze. They were trained on the entire double alternation sequence; *RRLL* or *LLRR*. The training was not given in four stages as it was with Group 1. Instead, the animals were trained in one stage which contained all four turns in the double alternation pattern.

At the end of the 85 training trials, all the animals (20) were given 10 test-trials, 5 trials a day. A test-trial is defined as eight unrestricted turns (all doors open) with reward at the end, regardless of the correctness or the incorrectness of turns.

Results. Following the same statistical procedures as the Gallup and Diamond study, a double alternation was defined as four turns, two in one direction followed by two in the opposite direction. In a given test-trial, where eight turns are made, there are five sets of four consecutive turns, or as they shall be labeled here, five quadruples. For example, with turns 1,2,3,4,5,6,7,8 the animal could double alternate on turns 1234; 2345; 3456; 4567; or 5678. With a choice of either left or right on each turn, and with eight consecutive turns, there are 256 possible combinations. Of these 256 combinations, there are 130 in which no quadruples can be scored correct; *e.g. RLRLRLRL RRLRLLRL.* There are 95 combinations in which one quadruple is correct; *e.g. LLRRLRLR;* there are 29 combinations in which two quadruples are correct; *e.g. LRRLLRL;* and there are two combinations in which three quadruples are correct; *e.g. RRLLRRLL* and *LLRRLLRR.* It is not possible to have four or five correct quadruples. The curve of expected frequencies approximates normality with a mean equal to 0.62, *SD* 0.07.[11]

---

[11] The author is indebted to the late Dr. Robert Beinert, Professor of Mathematics at Hobart College, Geneva, New York, for his untiring efforts in developing the statistical analyses and formulation.

The number of correct quadruples on each test-trial for animals in both groups was recorded and mean values were subjected to t-tests (Tables I and II). Each animal's total number of double alternations was also plotted against running times further to illustrate the vast differences between groups (Fig. 2). Results indicate that animals of Group 1 double alternated 104 times to 25 times for the animals in Group 2. T-tests clearly demonstrate that the *S*s of Group 1 significantly learned the double alternation pattern in the temporal maze ($p < 0.01$). The significance of the performance of Group 1 over that of Group 2 emphasizes the differences between the two training procedures utilized in this study.

TABLE I

CORRECT QUADRUPLES PER TRIAL AFTER TRAINING FOR GROUP 1 *S*S

Trial	Possible correct quadruples			
	0	1	2	3
I		7	2	1
II	2	7	1	
III	4	3	3	
IV		9	1	
V	3	4	3	
VI	3	4	3	
IIV	2	6	2	
VIII	3	5	2	
IX	1	8	1	
X	1	6	3	
Sum	19	59	21	1
Frequency	0	59	42	3

Frequency sum = 104; Mean = 1.04

TABLE II

CORRECT QUADRUPLES PER TRIAL AFTER TRAINING FOR GROUP 2 *S*S

Trial	Possible correct quadruples			
	0	1	2	3
I	7	3		
II	6	4		
III	7	2	1	
IV	8	2		
V	6	4		
VI	7	2	1	
VII	9	1		
VIII	7	3		
IX	10			
X	10			
Sum	77	21	2	0
Frequency	0	21	4	0

Frequency sum = 25; Mean = 0.25

Discussion. The results indicate that Group 1 did significantly better than Group 2 in double alternation performance. This upholds the theory expressed in the first hypothesis, *i.e.* that there will be a significant difference in the double alternation behavior of rats in a temporal maze that can be attributed to a specific training procedure. One reason for the difference demonstrated here might be found in a closer investigation of the point at which the animals were reinforced in their training sequences. The animals in Group 1 were reinforced 45 times after making right turns only, and 40 times after making right turns followed by left turns. The animals in Group 2 were reinforced 85 times after two right turns followed by two left turns.

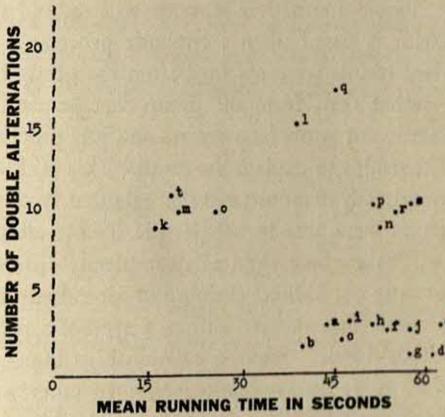


FIG. 2. ALTERATIONS AS CONTRASTED WITH RUNNING TIMES OF ALL Ss

In test-trials, animals of Group 2 tended to start off with left turns rather than right turns which served to impede them in the successful completion of the double alternation-pattern. Many of them started off with left turns and went left all eight times, omitting any right turns.

The animals in Group 1, however, tended to start off with right turns and then they completed the pattern with two left turns. We may conclude then, that if the animal is rewarded only after successful completion of a particular turn (under Group 2 procedures, a left turn), it would hinder complete learning, while if the animal is rewarded after different turns (Group 1) it tends to aid successful double alternation learning.

These two conditions may be viewed in the light of the Hullian concept involving delay in reinforcement after each turn on the first stage of learning. On the second stage, they received reinforcement after two turns, after three turns on the third stage, and after four turns on the fourth stage. Group 2 received reinforcement only after four turns on all 85 trials. Hull states that the greater the delay in reinforcement, the weaker will be the resulting reaction potential. In other words, the delay in reinforcement in animals of Group 2 could account for their poor learning. It would further explain why they persisted in making left turns. They were rewarded immediately after the left turns which were learned rather than the entire (RRLL) sequence.

In a discussion of the second hypothesis, that the training procedure that best demonstrates double alternation behavior will uphold the theory that this type of behavior is based upon a symbolic process, it would appear that because the best results were obtained from the group that learned in successive stages rather than from the group that learned in one stage, we would have to rule out symbolic processes and rely on chaining of stimulus response relationships to explain the results. This is, however, not the case, since we have already demonstrated that rats that have learned the sequence in four stages were able to transfer the double alternation pattern to another temporal maze which required a completely different set of muscular movements ruling out a direct chaining of kinesthetic responses. This suggests the hypothesis that the rats utilize a symbolic process in which they learn "substitute stimulus-response relationships that can be retained and later recalled to guide behavior selectively." In other words, although it appears that the animals learn the procedure by chaining responses, they seem to utilize a symbolic process in the actual performance of the double alternation pattern. Rather than learn a "cognitive map" of the physical maze, they learn instead a "cognitive pattern" that they can transfer to another temporal situation.

The results are interpreted to show that rats can learn to double alternate in a temporal maze, and that such learning is dependent upon substitute stimulus-response relationships rather than upon response chaining to kinesthetic stimuli. Correctness of a series of turns was scored in a way which permits partial correctness rather than the previously used, more stringent criterion of eight consecutive turns with no errors.

Summary. The literature on double alternation has suggested two major theoretical positions: Kinesthetic impulses arriving in the central nervous system from one response serve as cues for the next response in a chain; or an animal learns the double alternation pattern through the use of a symbolic process. Gallup and Diamond demonstrated significantly that rats can learn to double alternate and they suggested a number of questions. The present study was designed to investigate these problems and to tie in the specific training procedures with a single theoretical position. The major hypothesis is that the training procedure that best demonstrates double alternation will be applicable to symbolic process learning.

The statistical procedures utilized were defined in terms of quadruples and the number of correct quadruples on each test trial was recorded. Mean values were obtained and *t*-tests utilized. The results indicate the following: (1) The group that learned by chaining turns did significantly better than the group that learned by four-turn training, ($p < 0.01$); (2) the training procedure that best demonstrated double alernation was the chaining procedure which had been previously used in successful transfer of double alternation behavior, thereby demonstrating symbolic process learning.

EXPERIMENTAL CONSIDERATIONS CONCERNING CATEGORY AND MAGNITUDE SCALING

By M. B. DONNELLY and H. J. A. RIMOLDI, Loyola University
Chicago, Illinois

The purpose of this study is: (a) to investigate the values that subjects (*S*s) assign to a collection of stimuli presented singly and in all their possible paired combinations; and (b) to determine the effect of different scaling transformations on these values. We shall attempt, further, to explore experimentally issues concerning the relation between category- and magnitude-scaling. This article is part of a more extensive study from which several publications have already been presented.[1] More specifically, we were concerned with how *S*s will rate stimuli: (a) when asked to assign values to each member of the pair when the sum of the two values would be constant (constant sum method); (b) when giving values to each member of the pair without being concerned with a constant sum for each pair; and (c) when rating stimuli singly, using the successive category procedure.

Recently, Stevens (1966) summarized the literature pertinent to category and magnitude scaling.[2] The controversey originated in a study by Stevens in which he proposed a new psychophysical law and discussed the type of units used in Fechnerian and Thurstonian scaling as opposed to his magnitude scaling procedures.[3] According to Stevens, it is the *S*s' resolving power which defines the unit in the Fechner-Thurstone procedures, that is, the *S*s' variability or "confusion." Evidence for a concave downward relationship between magnitude- and category-data on prothetic continua for psychophysical data was presented by Stevens and Galanter (1957).[4] More recently, evidence for this relationship for non-metric stimuli has been presented by Ekman and Kunmapas studying handwriting;[5] by Engen and McBurney when scaling pleasantness of odors;[6] by Dudek when work-

---

* Received for publication March 31, 1967.

[1] J. R. Devane and H. J. A. Rimoldi, The determination of interval limits, Psychometric Laboratory Publication #16, Chicago: Loyola University, 1960, 1-4; The experimental study of the value of stimuli in combinations, Psychometric Laboratory Publication #20, Chicago: Loyola University, 1961, 1-9.

[2] S. S. Stevens, A metric for the social consensus, *Science*, 151, 1966, 530-541.

[3] Stevens, On the psychophysical law, *Psychol. Rev.*, 64, 1957, 153-179.

[4] S. S. Stevens and E. H. Galanter, Ratio scales and category scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1957, 377-411.

[5] Gösta Ekman and Teodor Künnapas, Measurement of aesthetic values by direct and indirect methods, *Scandinavian J. Psychol.*, 3, 1963, 33-39.

[6] Trygg Engen and D. H. McBurney, Magnitude and category scales of the pleasantness of odors, *J. exp. Psychol.*, 68, 1964, 435-440.

ing with favorableness of adverbs;[7] by Koh and Koh in a study of musical preferences; in normal and mentally retarded *Ss*;[8] and by Sellin and Wolfgang in appraising seriousness of criminal offenses.[9]

The downward concave relationship with respect to the magnitude-scale is usually logarithmic.[10] This relationship was interpreted by Stevens (1957, 1966) as due in the case of category scaling to the differential sensitivity of *Ss* throughout the continuum, in that *Ss* tend to be more sensitive at the lower end than at the upper end.[11]

Torgerson explains the differences between magnitude- and category-scaling in terms of the properties of the number systems of the two methods.[12] To quote him: "If we assign numbers so that a change from 10 to 11 pounds is the same size as a change from 1000 to 1001 pounds we get one type of scale [category equal interval-scale]. We get a second type of scale [magnitude-ratio scale] by assigning numbers so that the change from 10 to 11 is more equal to say a change from 1000 to 1100. These two ways of assigning values are logarithmically related. Thus the logarithmic relationship between magnitude scaling and category scaling may be partly explained by the different ways in which *Ss* assign numbers to stimuli.[13]

Method: (1) Subjects. The *Ss* were 122 men and women, students from several introductory psychology courses. Each *S* received a booklet including four sections. Fifteen value-statements[14] were presented to each *S* in the following manner and order:

Part *A* consisted of 105 pairs made by combining the 15 single statements in all possible pairs. Each *S* was instructed to assign any value between 0 and 100 to each member of the pair in accordance with his degree of agreement with each statement. *Ss* were specifically instructed to assign these values only after reading carefully each statement in the pair. They were told not to use the extreme values unless they felt complete agreement or disagreement with the statement.

Part *B* consisted of the same 105 pairs, but the *Ss* were instructed to assign to each statement in the pair any value between 0 and 100 provided that the sum of the values added up to 100. This is the constant sum method. Other instructions were similar to those in Part *A*.

In Part *C*, the *Ss* were asked to give a value from 0 to 100 to each one of the 105 pairs as such. The results of this study have been reported elsewhere by Rimoldi and Devane.[15]

---

[7] F. J. Dudek, A comparison of scale values for adverbs determined by the constant sum method and a successive intervals procedure, *Educ. Psychol. Measmt.*, 19, 1959, 439-448.

[8] S. D. Koh, Scaling musical preferences, *J. exp. Psychol.*, 70, 1965, 79-82; S. D. Koh and Tong-He Koh, Scaling of musical preferences by the mentally retarded, *Science*, 153, 1966, 432-434.

[9] Thorsten Sellin and M. E. Wolfgang, *The Measurement of Delinquency*, 1964, 260-261.

[10] Gösta Ekman and Lennart Sjöberg, Scaling, in P. R. Farnsworth (ed.), *Annual Review of Psychology*, 16, 1965, 464.

[11] Stevens, *opp. citt.*, 530, 154.

[12] W. S. Torgerson, Quantitative judgment scales, in Harold Gulliksen and Samual Messick (ed.), *Psychological Scaling: Theory and Applications*, 1960, 21-31.

[13] Torgerson, *op. cit.*, 25.

[14] These statements will be provided upon request.

[15] H. J. A. Rimoldi and J. R. Devane, Some considerations on scaling procedures, *Percept. mot. Skills*, 11, 1960, 207-213.

Part D consisted of the 15 single statements which the *S*s were to rate using the successive intervals method. The range of the categories went from —4 to +4.

Scoring procedures. Part A: Since every statement appeared in 14 different pairs, the 14 values assigned to each statement were averaged, obtaining the M_{Ai} values, where i represents a given statement ($i = 1$ to 15) and A indicates that these values were obtained from Part A.

Part B: The data from Part B were scaled using the constant sum method.[16] The values obtained from this procedure will be designated by S_{Bi} where B indicates Part B and i refers to a given statement. Also, the same averaging procedure used in Part A was employed with Part B. These values will be designated by M_{Bi}.

Part D: The 15 single statements were scaled using the procedure described by Rimoldi and Hormaeche.[17] The scale values thus obtained will be referred to as S_{Di}. In addition, a M_{Di} value was found for each statement by averaging the numerical value of the categories in which a given statement had been placed by all the subjects.

TABLE I

SCALE-VALUES FOR PARTS A, B AND D

Stimuli	M_{A_i}	M_{B_i}	S_{B_i}	M_{D_i}	S_{D_i}
1	37.6	40.7	.66	3.38	−.48
2	42.5	40.9	.67	3.95	−.29
3	33.1	37.0	.56	2.66	−.80
4	82.8	71.7	2.64	7.38	1.23
5	69.6	61.6	1.64	5.97	.79
6	46.2	44.2	.77	4.02	−.28
7	42.3	43.6	.75	3.89	−.27
8	41.8	43.3	.74	3.11	−.65
9	36.4	40.8	.66	3.21	−.59
10	41.8	42.3	.71	3.52	−.45
11	38.8	45.4	.82	3.10	−.75
12	31.6	35.8	.54	2.15	−.93
13	82.2	70.3	2.48	7.16	1.31
14	69.9	61.4	1.63	6.12	.85
15	82.5	71.0	2.56	7.24	1.32

Results. In Table I the values for all the scales are presented. Graphs were prepared by plotting the columns of Table I in all possible combinations. Only those graphs which are of special interest will be presented. Comparing the values of S_{Di} and M_{Bi} a near-perfect linear relationship was obtained (Fig. 1). Therefore for future comparisons only the S_{Di} values will be used. In contrast, the graph of S_{Bi} and M_{Bi} values indicated a perfect linear relationship between log S_{Bi} and M_{Bi} (Fig. 2) In Fig. 3, the graph for S_{Bi} and S_{Di} values is presented. Again, a logarithmic transformation of the S_{Bi} values gives a better linear fit. Figs. 4 and 5 show

---

[16] Torgerson, *Theory and Methods of Scaling,* 1958, 105-112.

[17] H. J. A. Rimoldi and Marcella Hormaeche, The law of comparative judgment in successive intervals and graphic rating methods, *Psychometrika,* 20, 1955, 307-318.

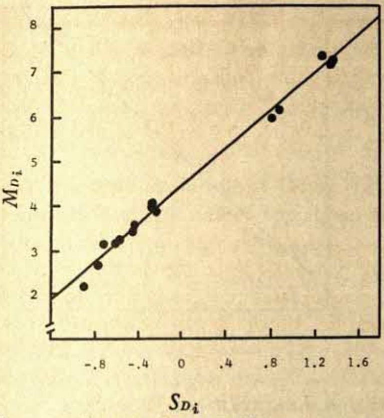


FIG. 1. LEAST SQUARE FIT OF THE SCALE-VALUES FOR SINGLE STIMULI (S_{Di}) AGAINST MEAN CATEGORY-RATING FOR THE SAME STIMULI USING THE SUCCESSIVE INTERVAL-PRESENTATION (M_{Di})

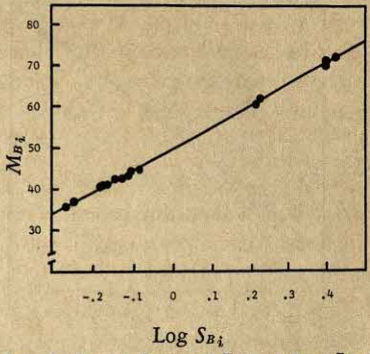


FIG. 2. LEAST SQUARE FIT OF THE LOG OF THE CONSTANT SUM METHOD (S_{Bi}) AGAINST MEAN VALUE FOR EACH STIMULUS DETERMINED BY AVERAGING THE NUMERICAL VALUES GIVEN TO EACH STIMULUS IN ALL THE PAIRS IN WHICH IT APPEARS USING THE CONSTANT SUM PRESENTATION (M_{Bi})

respectively that the M_{Ai} values are linearly related to the M_{Bi} values but logarithmically related to the S_{Bi} values. Finally, Fig. 6 and 7 indicate that M_{Ai} and S_{Di} are linearly related, while M_{Bi} and S_{Di} show a slight concave downwards effect.

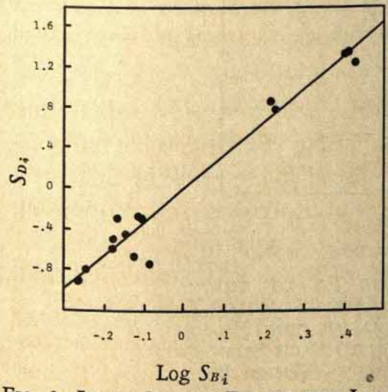


FIG. 3. LEAST SQUARE FIT OF THE LOG OF THE CONSTANT SUM METHOD (S_{Bi}) AGAINST SCALE VALUES OF THE SUCCESSIVE INTERVAL METHOD (S_{Di})

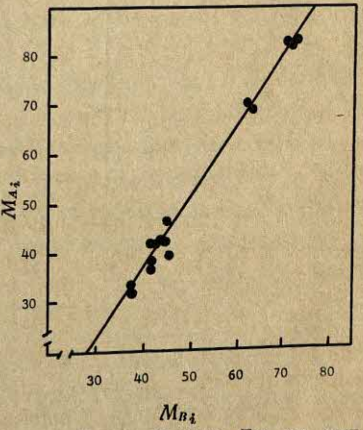


FIG. 4. LEAST SQUARE FIT OF THE MEAN NUMERICAL VALUE FOR EACH STIMULUS UNDER THE CONSTANT SUM RESTRICTION (M_{Bi}) AGAINST MEAN NUMERICAL VALUE FOR EACH STIMULUS WHEN NO CONSTANT SUM IS IMPOSED (M_{Ai})

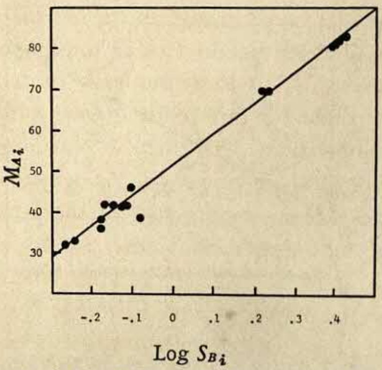


Fig. 5. Least Square Fit of the Log of the Constant Sum Method (S_{Bi}) Against Mean Numerical Value for Each Stimulus When No Constant Sum is Imposed (M_{Ai})

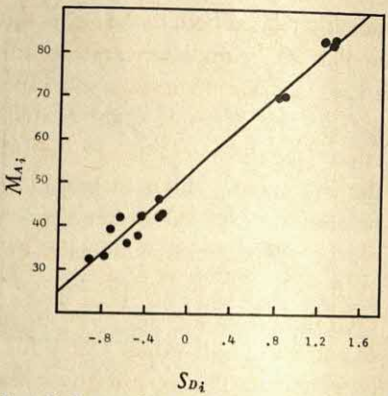


Fig. 6. Least Square Fit of the Scale Values of the Successive Interval Method (S_{Di}) Against Mean Numerical Values for Each Stimulus When No Constant Sum is Imposed (M_{Ai})

Discussion. In the first place, the usual linear relationship was found between the S_{Di} and M_{Di} values, which indicates that in all likelihood the unscaled values were normally distributed. Further, the M_{Ai} values are

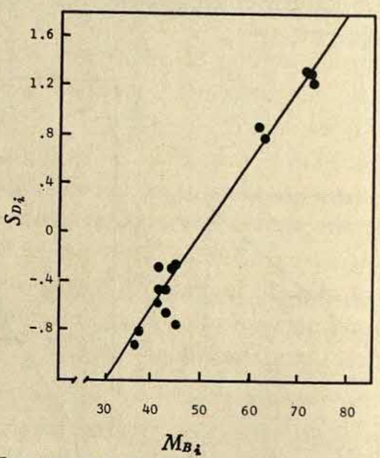


Fig. 7. Least Square Fit of the Mean Value for Each Stimulus Determined by Averaging the Numerical Value Given to Each Stimulus in All the Pairs in Which It Appears Using the Constant Sum Method (M_{Bi}) Against the Scale Values of the Successive Interval Method (S_{Di})

linearly related both to the S_{Di} and to the M_{Bi} values in spite of the fact that, in Part *D*, stimuli were rated individually and that in Part *B* the *S* was asked to make a constant sum. In contrast, M_{Ai} is logarithmically related to S_{Bi}. This strongly suggests that the reason for this concave downward effect (logarithmic) may be due to the operations performed in obtaining the S_{Bi} values, rather than to a psychological phenomenon. As a matter of fact, the S_{Bi} values are the averages (geometric means) of the ratio of the i^{th} stimulus with each one of the other stimuli; that is, $S_{Bi} = \sqrt[n]{R_{i1}\, R_{i2} \ldots R_{in}}$ where $R =$ a ratio, and $i = 1, 2, \ldots n$. Thus, whenever S_{Bi} values are compared with the *M* values for either Part *A* or *B* or *D* and with scale-values (S_{Ci}) for Part *D*, a nonlinear relationship should be obtained. This relationship should approach, within a margin of error, a logarithmic relationship. Notice that our results show exactly this.

Furthermore, the M_{Bi} values approach linearity when compared to the other *M* and S_{Ci} values with the exception of the S_{Bi} values, where the relationship is logarithmic. But the M_{Bi} values were obtained from the same raw data and thus from the same psychological process as the S_{Bi} values, the only difference being that the S_{Bi} values are expressed as ratios.

Notice that the relationship between M_{Ai} and M_{Bi}, as well as that between M_{Ai} and S_{Di}, is linear. Curiously, there is a slight concave downward effect between M_{Bi} and S_{Di}. Logically this should not occur, and therefore an explanation of the observation is not readily formed. Of course the number of categories used, the spacing of the stimuli and so forth could be used to explain it.

An interesting sidelight of this study is that the scale-values from Part *C*, where *S* judges the pairs as such, are linearly related to the sum of the values for each pair as obtained in Part *A*, as well as to the combined values from Part *D*, when using a method developed by Rimoldi.[18]

All these findings strongly suggest that an explanation for the concave downward relationship between category and magnitude scales may be due to the mathematical transformations used in the different scaling methods.[19]

---

[18] H. J. A. Rimoldi, Predictions of scale-values for combined stimuli, *Brit. J. Stat. Psychol.*, 9, 1936, 29-40.

[19] The authors are deeply indebted to Mrs. M. Aghi, Mrs. P. Dore, Mr. K. Vander Woude, and Mr. T. Chlapecka for their assistance in this study.

EFFECT OF SPEED OF MOVEMENT ON TACTUAL-KINESTHETIC PERCEPTION OF EXTENT

By SEYMOUR WAPNER, JOSEPH WEINBERG, J. A. GLICK, and GEORGE RAND, Clark University

The analysis of perception of extent in the visual modality has been largely restricted to a static, geometric characterization in terms of the dependency of retinal size on the physical size and distance of an object from the retinal plane ("size-distance invariance hypothesis").[1] In contrast, analysis of perception of extent in the tactual-kinesthetic modality provides the basis for a dynamic analysis of the relation between quality of organismic activity (*e.g.*, speed of movement) involved in perception, and object properties (*e.g.* size, extent) which are perceived through such activity. It is assumed that variation in quality of activity (*e.g.* fast vs. slow rate of traversing a given physical extent) makes for a difference in perceived extent.

Such an analysis was suggested by Brown a long time ago in his classic treatment of perception of extension, where he attempted to show that temporal succession of muscular activity is an essential feature of extensionality.[2] He suggested, without supporting data, that the slower the speed of traversing the extent, the greater the extent appears. This serves as the hypothesis for the present study.

Method. By means of a specially constructed apparatus, *S*'s finger and outstretched arm was moved passively, at chest level, from left to right (right to left) in a horizontal line over an extent of 40 cm. in his fronto-parallel plane. The beginning and end points of this extent were fixed symmetrically with respect to

---

* Received for publication February 22, 1965. This investigation was supported by research grant MH 00348 from the National Institute of Mental Health, Public Health Service. We wish to express our appreciation to Miss Jill Rierdan for assisting in collecting the data. The second author (Joseph Weinberg) is currently located at the Institute of Physical Medicine and Rehabilitation, New York, N. Y.; the third author (J. A. Glick) is currently located at Yale University.

[1] William Epstein, John Park, and Albert Casey, The current status of the size-distance hypotheses, *Psychol. Bull.*, 58, 1961, 491-514.

[2] Thomas Brown, *Lectures on the Philosophy of the Human Mind*, 16th ed., Vol. 1, 1846, 533-548. These experiments were conducted prior to the authors' contact with Brown's work, *ibid.*, which came to their attention when an excerpt was reprinted recently in William Dember, *Visual Perception: The Nineteenth Century*, 1964, 102-113. See also C. O. Weber, The properties of space and time in kinaesthetic fields of force, this JOURNAL, 38, 1927, 597-606; and Theodora M. Abel, A comparison of tactual-kinesthetic and visual perceptions of extent among adults, children, and subnormals, this JOURNAL, 48, 1936, 269-296.

S's objective median plane. In the course of traversing the total extent *S*'s finger touched an intermediate point ("center" marker), which divided the total extent into two parts: extent of first phase (from beginning point to center marker); and extent of second phase (from center marker to end point). The apparatus was so arranged that the speed of each "phase" could be independently adjusted as either slow or fast. The task for *S* consisted in establishing equality between the first extent traversed at a particular speed (fast, slow) and the second extent traversed at the same or different speed.

Apparatus. The apparatus, schematized in Fig. 1, has a finger plate (A) which moves from left to right (right to left) from a beginning point I (III) past an intermediate point II to a terminal point III (I) along a horizontal distance of

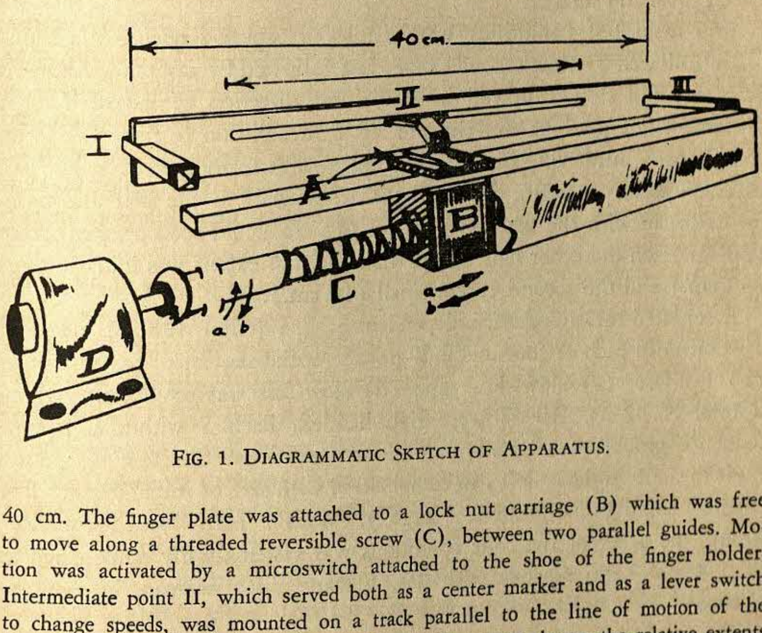


FIG. 1. DIAGRAMMATIC SKETCH OF APPARATUS.

40 cm. The finger plate was attached to a lock nut carriage (B) which was free to move along a threaded reversible screw (C), between two parallel guides. Motion was activated by a microswitch attached to the shoe of the finger holder. Intermediate point II, which served both as a center marker and as a lever switch to change speeds, was mounted on a track parallel to the line of motion of the finger, and its position could be varied in small steps to change the relative extents of the first and second phases of movement. The screw controlling speed of linear motion was moved by a DC motor (D) (115 volts; 1/15th hp). The center marker (II) could be set to shift from slow to fast speed, fast to slow speed, or to maintain movement at the same speed, slow, or fast. Two speeds of linear movement, calibrated by motion picture photography, were employed, viz., a slow speed (4 cm./sec.) and a fast speed (5 cm./sec.).

During the experiment, *S* was blindfolded and his finger, placed on the finger holder, was moved over the horizontal 40 cm. distance. At a particular point in the excursion *S* touched the center marker, which delimited two extents. *S* was required to tell *E* whether the first extent was shorter, equal to, or longer than the second. Care was taken to insure that *S* understood the instructions by having him make a judgment on a visual basis before blindfolding him for the actual trials. At the beginning of a particular trial, the first extent traversed was small (large) compared with the second extent. Following that trial, the difference in extent was reduced by 2 cm. and the next was carried out. This procedure was continued until *S* shifted his judgment from shorter (longer) to longer (shorter) for three consecutive trials. Essentially then, a procedure involving a variation of the method of limits was employed. The Point of Subjective Equality (*PSE*) was determined from the transition points between larger and equal, and between equal and smaller.

Two main test conditions were employed: speed-sequence (fast first phase followed by slow second phase; fast followed by fast; slow followed by slow; and slow followed by fast); and direction of motion (finger movement left to right and right to left). In addition to these two main conditions, a third, viz., size of initial extent, was introduced as part of the psychophysical procedure described earlier: on half of the trials the judgments began with the first extent small (10 cm.) and second extent large (30 cm.); on the other half of the trials the first extent was initially large (30 cm.) and the second extent small (10 cm.).

A repeated measurement, independent groups design with 8 *S*s (4 men; 4 women) in each of four methods-groups (speed-sequence: (a) fast-slow; (b) fast-fast; (c) slow-slow; and (d) slow-fast was employed, making a total of 32 *S*s. All *S*s were right-handed. Each *S* within a "speed-sequence group" was tested under four conditions: direction of motion (left to right; right to left) in combination with size of initial extent (first extent "small"; first extent "large"). Within each "speed-sequence" group, the order of the four conditions was systematically randomized as in a 4 × 4 Latin square, with 2*S*s, 1 man and 1 woman, in each sequence.

The measure employed in the analysis consisted of the physical distance —computed from PSE—from the beginning point to the center marker (*apparent extent of first phase*).

Results. Before considering the efficacy of the main experimental variable, the results were examined to determine whether there were over-all differences depending upon temporal order (time-error). For this purpose, the means for the "slow-slow" and "fast-fast" speed-sequence groups were examined. Since the measure employed is physical extent of first phase (extent of second phase is the remainder of the 40 cm. distance), lack of effect of temporal order would be reflected in values approximating

20 cm. Examination of the findings for the "fast-fast" and "slow-slow" speed-sequence groups (21.16 cm. and 21.61 cm. respectively) shows that in each instance, the physical extent for the first phase was made larger than that for the second phase in order to experience the total extent as divided into equal halves. When these mean values were evaluated against the hypothesis that the population mean is 20 cm., they were both significant ($p < 0.05$). Thus, the physical extent of the first phase must be made significantly greater than the physical extent of the second phase in order for the two extents to appear equal. Stated another way, when traversing a total extent of 40 cm. the first half appears smaller than the second half.

The test of the main hypothesis involves comparison of the mean apparent extent for the "fast-slow" and "slow-fast" speed-sequence groups. Independent of other factors, the hypothesis calls for the first phase of the "fast-slow" speed-sequence group to be greater than 20 cm.; and the first phase of the "slow-fast" speed-sequence group to be smaller than 20 cm. Since, as suggested above, temporal order has an over-all effect, this difference should hold only in relative terms, *i.e.,* for the measure employed—physical extent of first phase—the "fast-slow" speed-sequence group should yield a *relatively* larger value than the "slow-fast" sequence group. Analysis of variance revealed that the speed-sequence variable was significant ($F = 3.08$, 3 and 24 df). The mean for the "fast-slow" group (21.39 cm.) is significantly larger than the mean for the "slow-fast" sequence group (19.97 cm.).[3] A less conclusive secondary finding pertained to differences due to direction of arm movement. This bears on the problem of differences in tactual-kinesthetic extent resulting from movement of the arm in contralateral or ipsilateral space and should be controlled in this type of experimentation.

Discussion. In keeping with Brown's observations, the major finding of this study is that a physical extent traced at a given speed must be made relatively long in order to appear equal to the same extent traced at a slower speed. This finding has significance insofar as it augments the variety of situations in which an interrelation has been demonstrated between temporal features of stimulation and perceived spatial extent. Such space-time interactions, first demonstrated by Benussi[4] and Gelb,[5] and more recently by others,[6] characteristically involve successive point stimulation, and are of two types: the *tau*-phenomenon (the shorter the physical time between successive spatially distributed flashes, the shorter the perceived distance between them) and the *kappa*-phenomenon (the shorter the physical extent between successive flashes, the shorter the apparent time). In our study analogous results were obtained with continuous motion.[7]

---

[3] A preliminary experiment conducted by Joseph Weinberg (Effect of speed of passive finger movement on perception of extensionality, Unpublished Master's thesis, Clark University, 1963) produced essentially the same findings with an independent group of 32 *Ss*.

[4] Vittorio Benussi, *Psychologie der Zeitauffassung*, 1913, 349.

[5] This has also been demonstrated for the tactual modality of Adhemar Gelb,

Taken together, these findings of space-time interaction provide a complexity which is outside the scope of traditional accounts of perception of extent: On the one hand, the traditional approach has employed a spatial model in which perceived extent is assumed to be based primarily on the retinal distribution of points of stimulation with an assumed correspondence between proximal stimulus and perceptual experience. The inadequacy of such a retinal model is particularly evident in its inability to deal with the space-time interaction, namely, the lack of functional equivalence between identical sets of spatial stimuli where temporal relations are different. An alternative way to handle the space-time problem—and to retain nonetheless the notion of one-to-one relation between proximal stimulus and perception—would be to redefine the nature of the proximal stimulus as a "spatio-temporal entity." However, this alone is not satisfactory since it raises the theoretical issue of how a single proximal stimulus (a spatio-temporal entity) may function alternatively as either the occasion for perception of a spatial property or as the occasion for perception of a temporal property.

In our own view, perception of extent is mediated by common organismic processes which provide the basis for the interaction of spatial and temporal characteristics. While we cannot as yet clearly define the nature of this common process, the present study, taken together with others on space-time interaction, indicates that this process must be conceived broadly

---

Versuche auf dem Gebiete der Zeit-und Raumanschaung, *Bericht uber der VI. Kongress fur Experimentelle Psychologie,* 1914, 36-42.

[6] Harry Helson, The *tau* effect.—An example of psychological relativity, *Science,* 71, 1930, 536-537; Harry Helson and S. M. King, The *tau* effect: An example of psychological relativity, *J. exp. Psychol.,* 14, 1931, 202-217; P. E. Comalli, Jr., The effect of time on distance perception, Unpublished Master's thesis, Clark University, 1951; John Cohen, C. E. M. Hansel, and J. D. Sylvester, A new phenomenon in time judgment, *Nature,* 172, 1953, 901; Cohen, Hansel, and Sylvester, *op. cit.,* Interdependence of temporal and auditory judgments, *Nature,* 174, 1954, 642-646; Cohen, Hansel, and Sylvester, *op. cit.,* Interdependence in judgments of space, time and movement, *Acta Psychologica,* 11, 1955, 360-372; D. R. Price-Williams, The Kappa Effect, *Nature,* 173, 1954, 363-364. Some of these and other studies are presented in Paul Fraisse, *The Psychology of Time,* 1963, 136-137.

[7] It should be mentioned that continuous movement-effects were also found by G. S. Hall and H. H. Donaldson, Motor sensations on the skin, *Mind,* 1885, 557-572.

enough to allow for the possibility of interactions involving both successive point stimulation and continuous movement.[8]

SUMMARY

Variation in speed of tactual-kinaesthetic tracing a given physical extent significantly affects perception of that extent: with relatively faster (slower) speed a given extent is perceived as relatively shorter (longer).

---

[8] Since analogous effects have been demonstrated for distances measured in miles and for times measured in hours by John Cohen and Peter Cooper, New phenomena in apparent duration, distance and speed, *Nature;* 196, 1962, 1233-1234, it would be ideal if the conceptualization could encompass this added dimension of diversity.

"CLOSURE" IN A GAPPED CIRCLE FIGURE

By DAVID S. HOLMES, University of Texas

The Gestalt principle of "closure" predicts that in time the memory-trace for a circle with a gap will undergo an autonomous change in which the gap will become smaller.[1] That is, it is hypothesized that because the gapped circle figure is not in a state of configural or conceptual "balance," in time, the memory-trace for the figure will undergo a change and the disturbed figure will become a "better" or more balanced figure. A second hypothesis predicts that the memory-trace can be influenced by traces from intervening patterns.

Hebb and Foord reviewed the research on closure published prior to 1945 and found that, while there was unequivocal support for the hypothesis that "external" factors could affect the trace, because of methodological problems, no conclusions could be drawn concerning the hypothesized autonomous changes in the memory-trace.[2] More recently, Riley reviewed the research and again pointed out the numerous methodological problems found in the literature. Riley concluded by asking "whether the Gestalt hypothesis is even testable—at least by the methods which have been employed."[3] It is clear from previous research and criticism that in such a test the recognitive method should be used rather than the reproductive method, thus eliminating problems related to drawing ability and the subjective judgments of the drawings.[4] Further, different *S*s should be used to test recognition after different time-intervals rather than repeated testing of the same *S*s.[5] In addition, if an autonomous change in the memory-trace is to be assessed, words that are descriptive of the stimulus must not be used in the instructions to *S*s, since these may serve as labels which would influence the recall of the stimulus.[6] Finally, from the data of Carlson and Duncan, it appears that *S*s should not be given a series of similar stimulus-figures from which to identify the original figure shown to them because the nature of the figures which precede the original figure in the test-series will affect the *S*s' responses.[7]

---

[1] Kurt Koffka, *Principles of Gestalt Psychology,* 1935, 545.

[2] Donald O. Hebb and E. N. Foord, Errors of visual recognition and the nature of trace, *J. exp. Psychol.,* 35, 1945, 335-348.

[3] Donald A. Riley, Memory for form, in Leo Postman (ed.), *Psychology in the Making,* 1963, 402-465.

[4] Nelson G. Hanawalt, Memory trace for figures in recall and recognition, *Arch. Psychol.,* 31, 1937, (No. 216), 1-89; O. L. Zangwill, An investigation of the relationship between the process of reproducing and recognizing simple figures with special reference to Koffka's trace theory, *Brit. J. Psychol.,* 27, 1937, 250-276.

[5] Robert S. Woodworth, *Experimental Psychology,* 1938, 69-91; Hanawalt, *op. cit.,* 1-89.

[6] Leonard Carmichael, H. P. Hogan, and A. A. Walter, An experimental study of the effect of language on the reproduction of visually perceived form, *J. exp. Psychol.,* 15, 1932, 73-88; Riley, *op. cit.,* 402-463.

[7] J. B. Carlson and Carl P. Duncan, A study of autonomous change in the memory trace by the method of recognition, this JOURNAL, 68, 1955, 280-284.

The purpose of the present experiment was to assess the predicted autonomous change in the memory-trace of a circle with a gap. The broken circle was used because numerous other researchers have employed it and because the Gestalt prediction with regard to the nature of the change is clear. In an attempt to eliminate the confounds of previous research, different *S*s made judgments after different intervals of time, no words which could be used as labels were used in the instructions, and a type of recognitive measurement was used which did not involve the serial observation of numerous patterns in the search for the original stimulus.

Method: (1) *Subjects.* The *S*s were 768 students in introductory psychology classes at Northern Illinois University. Each of nine classes was tested as a group in the same classroom. Conditions were randomly assigned to classes. No indication of the purpose of the experiment was given, and the classwork had not yet touched upon either Gestalt theory or perception.

Diagrams and conditions. The original pattern for all nine conditions was a circle with a gap of 30° which was projected onto an elevated glass-beaded screen. The test-situation occurred after either 30 sec., 30 min., or 7 days. In the tests, two patterns were simultaneously projected upon the screen. These were identical to the first pattern except that they had gaps of 20° and 40°, deviations of 10° from the original pattern. The assignment of figures to the right and left sides was done randomly. Immediately preceding the presentation of the two test-patterns, one of the three groups in each time-condition viewed a circle with a 60° gap (twice the size of the gap in the original figure) for 15 sec. The second group viewed a complete circle for 15 sec. These were the groups used to study the effect of intervening patterns. A third group, used to test for the autonomous change, was not presented with any intervening patterns.

Procedure. At the beginning of the class period, the instructor introduced *E* to the class, asked the students to cooperate in a short experiment, and then left the room. The lights were turned off and *E* gave the following directions:

Instructions: I shall shortly project a figure on the screen and ask you to observe it. At some later time, I will show you a number of other figures and ask you to respond to them. For now, please just observe the figure shown to you. Here is the figure.

At that point, the figure with the 30° gap was projected for 15 sec. If the testing was not to take place immediately, *i.e.* under the 30-min. or 7-day condition, *E* left the room and classwork began. To initiate the testing after the appropriate interval for the intervening pattern conditions, *E* said:

Instructions (continued): Some time ago, you observed a figure that was projected on this screen. I wish you now to decide which of the figures I shall show you is identical to, or more closely resembles, the original figure shown to you. Please do not confer with those around you concerning your choice. First, I will project a test pattern to check the focus. [The intervening pattern—either the 60° gapped circle or the complete circle—was then projected for 15 sec. Following this, the 20° and 40° gapped circles were simultaneously projected for 15 sec. and the *S*s were asked to make their responses.] Please make your choice from between these figures.

The directions and procedures for the no-intervening pattern condition were the same, except that the comments on, and the presentation of, the "test pattern for focusing" were omitted. The *S*s recorded their choices on a standard form. At no time were the words "circle" or "gap" used; nor was it ever pointed out that this was a test of memory.

Results and discussion. Analysis of the data indicated the opposite of what the Gestalt closure-hypothesis would predict. Chi-square analyses of the proportion of *S*s opening the gap (*i.e.* selecting the figure with the 40° gap as identical to, or most closely resembling, the original stimulus) vs. the proportion of *S*s closing the gap indicated a strong over-all tendency toward *opening* rather than *closing*. In six of the nine conditions, as an inspection of Table I shows, significantly more *S*s recalled the gap as being larger. This was the case despite the fact that in some conditions the effect of the intervening stimuli was, according to the theory, working against the opening of the gap. It is interesting to note that, although it was rarely specifically pointed out in the reports, in a number of earlier studies which employed similar small-gapped circles, evidence of opening rather than of closing had been found.[8] It should be noted that in the one case in which a tendency toward closing was found,[9] the stimulus was considerably

TABLE I

DISTRIBUTION OF *S*S OPENING OR CLOSING THE GAPPED CIRCLE

	Intervening		
Time	Nothing	Circle	60° Pattern
30 sec.	Opened=38 (60%) Closed=25 $\chi^2=2.68$ ($p<0.10$)	Opened=45 (46%) Closed=53 $\chi^2=0.65$ ($p<0.50$)	Opened=44 (61%) Closed=27 $\chi^2=4.04$ ($p<0.05$)
30 min.	Opened=59 (66%) Closed=31 $\chi^2=8.71$ ($p<0.01$)	Opened=53 (56%) Closed=40 $\chi^2=0.91$ ($p<0.50$)	Opened=67 (62%) Closed=41 $\chi^2=6.26$ ($p<0.02$)
7 days	Opened=49 (65%) Closed=26 $\chi^2=7.04$ ($p<0.01$)	Opened=59 (74%) Closed=21 $\chi^2=18.05$ ($p<0.01$)	Opened=60 (67%) Closed=30 $\chi^2=10.00$ ($p<0.01$)

In all cases, $df=1$

---

[8] Hebb and Foord, *op. cit.*, 335-348; Carlson and Duncan, *op. cit.*, 280-284; Carl Duncan, Controlled fixation of the stimulus-figure in a study of autonomous change in the memory-trace, this JOURNAL, 73, 1960, 115-120; James C. Crumbaugh, Temporal changes in the memory of visually perceived form, this JOURNAL, 67, 1954, 647-658; Edward L. Walker and Joseph Veroff, Changes in the memory-trace for perceived forms with successive reproductions, this JOURNAL, 69, 1956, 395-402; Lawrence Karlin and Glen Brennan, Memory for visual figures by the method of identical stimuli, this JOURNAL, 70, 1957, 284-252.

[9] Walker and Veroff, *op. cit.*, 395-402.

different, *i.e.* the gap was at least twice the size of that used in other investigations. The opening of the gap, as found in the present and previous studies, might be attributed to "emphasis" or "sharpening," which refers to the tendency to exaggerate any unusual or distinctive feature because it is the only irregularity in an otherwise regular figure.[10]

The Gestalt theory also predicts that the memory-trace will go through a continuous, progressive, autonomous change. The present data afford no support for this prediction; for in the no-intervening pattern-condition, the proportion of *S*s opening the gap after 30 sec. was not significantly different from that of *S*s opening after 7 days. This would suggest that what change does take place occurs within the first few seconds after the trace has been laid down and then, for the most part, stops.

With regard to the influence of the intervening patterns on the trace, when the proportion of *S*s opening the gap in the 60° intervening pattern-condition was compared to the proporton of *S*s opening in the no-intervening pattern-condition, there was no significant difference after any time interval ($z = 0.19$, 0.51, and 0.78). On the other hand, the intervening circle did tend to increase the proportion of *S*s closing the gap, but this occurred only in the 30-sec. condition ($z = 1.79$, $p = 0.07$). While limited to only one of the recall periods, this effect was in the direction which would be predicted by Gestalt theory and previous data. That is, seeing a completely closed figure before the recall test resulted in a greater tendency for *S*s to close the gap in the test-figure. It might be suggested that the general lack of effects of intervening patterns in the present study, although found by other researchers,[11] is due to the fact that only one pattern was presented; that it was presented for only a short period of time, and that the *S*s did not carefully study it since it was only being presented to "check the focus" of the projection. In Carlson and Duncan's study, in which the effect was found, the *S*s saw a *series* of similar patterns, which they studied carefully because they were searching for the original stimulus.[12] It may very well be that the amount of attention *S*s devote to the intervening pattern is important.

Summary. The present research tested the Gestalt prediction that a circle with a gap would be recalled after a period of time as having had a smaller gap, *i.e.* the prediction of "closure" in the memory-trace. Using different groups of *S*s, *E* tested recall after 30 sec., 30 min., and 7 days, employing the method of recognition. The results indicated a significant opening of the gap rather than a closing, a finding contrary to the prediction of Gestalt theory. Further, there was no evidence of a progressive change in the memory-trace. Finally, in this investigation, intervening patterns had relatively little effect opon the memory-trace, but this may have been due to the short exposure of the intervening patterns.

---

[10] E. V. Fehrer, An investigation of the learning of visually perceived forms, this JOURNAL, 47, 1935, 187-221; James J. Gibson, The reproduction of visually perceived forms, *J. exp. Psychol.*, 12, 1929, 1-3; Friedrich Wulf, Ueber die Veränderung von Vorstellungen (Gedächtnis und Gestalt), *Psychol. Forsch.* 1, 1922, 333-373.

[11] Carmichael, Hogan, and Walter, *op. cit.*, 73-86; Riley, *op. cit.*, 402-463.

[12] *Ibid.*

APPARATUS

BONE-CONDUCTING MICROPHONES

By DONALD P. HAYES and LEO MELTZER, Cornell University

The faithful recording of speech in the presence of background noise is difficult to attain. The ability to make such recordings would be very useful to social psychologists, to linguists, therapists, and to mass-media specialists, in their transcriptions of the interactions of small groups, in their measurements of the temporal features of interaction, in making precise recordings of phonemes in natural conversation, in analyzing interviews between patient and therapist, and in broadcasting communications from areas of high noise-level such as conventions or sporting events. This paper describes a method of making such recordings which is superior, as we believe, to any at present available.

Problem. The major limitation to faithful recordings of speech in the presence of background noise is the sensitivity of conventional microphones to *all* noise in the vicinity of the speaker. Highly directional and 'close-talking' microphones have been used to deal with high noise-to-signal problems. The highly directional, fixed-position microphone is an unsatisfactory solution for the following reasons: (1) amplitude readings are adversely affected by such body-movements as tilting back in chairs or moving the head down or to the sides; (2) sounds originating or reflected from behind the speaker are not excluded; and (3) these microphones, while more sensitivie to front- than to side- or back-originating sounds, do *not,* in fact, eliminate sounds from these other sources.

'Close-talking' microphones of high quality are sometimes used in these situations. They are generally hand-held, as close as ¼-in. from the lip, or are boom mounted from a headset. Neither arrangement is well suited for 'normal' conversations, however useful they might be for other purposes. Experimental *S*s seated face-to-face cannot be counted upon to raise the microphones to their lips each time they wish to speak when the conversation becomes quite animated. What is more, even when mounted on a headset, 'close-talking' microphones do not succeed in eliminating normal voice-levels at 6 ft. away, at either side, or at 180° from the microphone.

There are several types of mechanical transducers which produce intelligible audio signals. Their advantages lie in being less responsive to sound sources other than those to which they are attached. A 'tooth' microphone is reported to be in use for communications purposes in the Air Force.[1] This is actually a denture which is unresponsive to loud external sounds such as the scream of a jet engine. The fact that it is a denture makes its use in most experimental research impractical.

'Throat'-microphones have been used with high levels of background noise. There are, however, several limitations to their use: (1) intelligibility is restricted because the higher frequencies of speech are generated in the oral cavity with the help of the tongue rather than in the throat to which the microphone is attached; (2) it is difficult to obtain comparability of microphone attachments to the throat from time to time or from person to person; and (3) while continuous pressure of the microphone against the throat improves intelligibility, continuous pressure is unquestionably uncomfortable and, for some, intolerable.

Solution: The Bone-Conductance Microphone. For over one year, we have been using a bone-conductance microphone attached to the forehead by means of a headband to record individual speech with good-to-excellent rejection of all other noise sources normally detected by conventional microphones (see Fig. 1). When people speak, vibrations prodoced by the sounds are carried through the cartilage to the forehead as well as the rest of the skull. Vibrations detected at the forehead have been shown to be more like those of sounds picked up at the mouth than at any other position on the skull or throat.[2] The transducer we use is a Sonotone 700-3001 Bone Oscillator having a 1000-cps impedance of 1000 Ω. It was made for

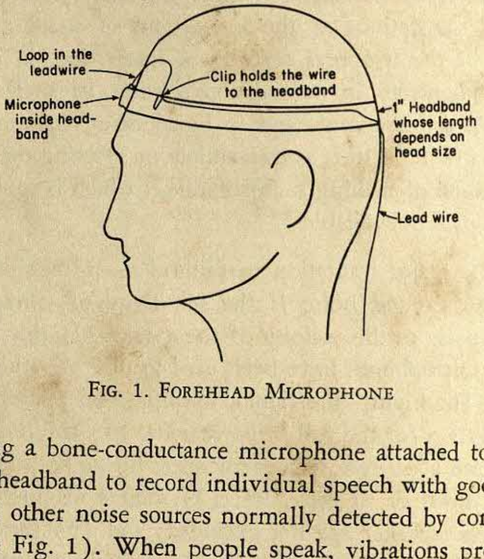

Loop in the leadwire
Microphone inside head-band
Clip holds the wire to the headband
1" Headband whose length depends on head size
Lead wire

FIG. 1. FOREHEAD MICROPHONE

---

[1] Personal communication from Mr. Veneklasen, President of Western Electro-Acoustic Laboratory, Inc., 2222 So. Barrington Ave., Los Angeles, Calif.

[2] K. K. Neely and S. E. Forshaw, Speaking and listening through the head: I. The intelligibility of speech recorded in quiet at different positions on the head and throat, *J. Auditory Research,* 5, 1965, 151-158; J. C. Snidecor, Irving Rehman, and D. D. Washburn, Speech pickup by contact microphone at head and neck positions, *J. speech hear. Res.,* 2, 1959, 277-281.

certain hearing-aid applications to vibrate the mastoid region, not to detect vibrations as in our case. A 1-in. headband holds and conceals the microphone and some of the wire which passes over the shoulder to the floor behind the person. The band's length is dependent on the *S*'s head size to ensure comparable microphone pressure on the forehead. The pressure on the forehead should be between 0.2 and 0.3-lb. per sq. in. None of the several hundred students tested has complained of discomfort while wearing this unit even for prolonged periods. The small shielded wire leaning from the microphone is connected to a coaxial cable at a distance of approximately 5 ft. The wire should be so clipped to the headband with some slackness between the microphone and the clip that a person's movements will not put any pressure upon his microphone. The terminal BNC connectors are so arranged that the signal can be carried either to the preamplifiers of an audio-tape recorder or directly to the preamplifiers of an analog-to-digital converter.

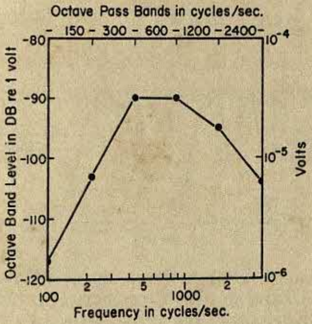


FIG. 2. SIGNAL CHARACTERISTICS OF SONOTONE 700-3001 MICROPHONE

Add —5 db. for raised voice-level. Levels of peak speech are —10 to —15 db. above long-term average. Measurements made by the Western Electro-Acoustic Laboratory, Inc., Los Angeles, California

Performance. The performance of this recording system warrants a detailed comparison with conventional microphones. The frequency response of this microphone is shown in Fig. 2. The two most notable features of this figure are: (1) the usable frequency range is approximately 225-2500 ~ per sec. without further electronic equalization to compensate for the weaker low and high frequency-signals; and (2) these microphones are detecting very low levels of sound pressure from the forehead (as little as 1μv. in normal conversation). This small output requires more than the usual pre-amplification to generate adequate signal-levels for most audio-tape recorders.

A direct comparison of a good quality conventional microphone (Shure 5785) with the forehead-unit is shown in Fig. 3. A short passage ("Now is the time (hesitation) for all good men . . .") was spoken while the speaker wore this unit. This passage was also detected by the conventional air-transmission microphone placed 3 ft. from the speaker's mouth. The simultaneous signals were directed to two previously balanced pre-amplifiers and thence to an analog-to-digital converter which sampled each channel at 33 times/sec using a time constant of 0.10 sec. The intensity of the voice at each sample is plotted in Fig. 3 along the vertical and time is plotted along the horizontal axis. The two curves (representing the sound intensities as read by the forehead and conventional microphones) are similar, both in terms of the points where the intensity rose or fell and in the durations of the speech and silence periods. The small divergence between the two curves at high intensity levels is only slightly larger than what is found when two identical conventional microphones are used to record the same sound.

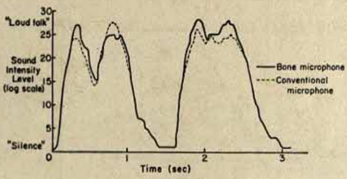


Fig. 3. Direct Comparison of the Forehead and Conventional Microphone on the Passage of *"Now is the time"* (hesitation) *"for all good men"* (Time-constant 0.10 sec. and sampling rate 33/sec. per data-source.)

A second point in comparing forehead with conventional microphones concerns the former's sensitivity to low intensity signal levels originating with the speaker. We have found that the forehead units are superior in detecting low level whispers and sounds such as "uh huh." These sounds can be quite important in affecting the course of interaction in groups, as Matarazzo and others have shown that the average duration of a speaker's utterance can be doubled or halved in contrast to his normal utterance length simply through the presence or absence and patterning of this "uh huh" sound on the "listener's" part.[3] Since the semantic content of speech by no means exhausts the significant information communicated in face-to-face interaction, there is ample reason to record as much of the interaction as is possible, whether the *S*s recognize that they "pay attention" to these sounds or not.

A final test of the forehead-units relates to their usefulness in eliminating signals other than those produced by the speaker; such as other's voices, footsteps, door closings and the like. Fig. 4 shows the intensive readings from an attached forehead-microphone while another person spoke the previously described passage at normal voice-level from a distance of 6 ft. The forehead-unit was unresponsive to this "noise" while the conventional microphone, likewise at 6 ft. from the speaker, produced a characteristic intensity-wave associated with this passage. It is this feature of the forehead-microphone which permits the unique identification of sound produced by the speaker and reduces the uncertainty in identifying who said what, when, and how.

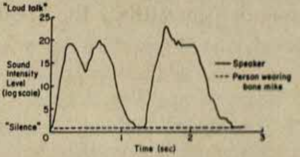


FIG. 4. NOISE REJECTION CAPABILITY OF THE FOREHEAD MICROPHONE
(Time-constant 0.10 sec. and sampling rate 33/sec. per data-source.)

Limitations in the use of bone-conductance microphones. As is so often the case, advantages of a piece of apparatus are offset by some disadvantages. These include: (a) surreptitious recording is precluded since this recording method requires the *S*'s approval and coöperation; (b) fidelity of recording is not high when compared to conventional microphones; (c) the *S*s cannot chew gum, gnaw on pencils, scratch their heads or fiddle with the microphone since all these acts produce signals which can be mistaken for speech; and finally, (d) as yet, we do not know the effects of skull thicknesses or clogged nasal passages on intensity-readings at the forehead. These limitations may be trivial or crucial, depending on the nature of the researcher's problem. Aside from one student who sincerely believed he would be given electro-shock through the microphone, the *S*s have graciously accepted the units as part of a laborabory environment.

---

[3] J. D. Matarazzo, George Saslow, Arthur Wiens, M. Weitman and B. Allen, Interview head nodding and interviewee speech durations, *Psychotherapy*, 1, 1964, 54-63.

Summary. The paper describes an audio recording system for those special applications where it is important to exclude all noises other than the vocal sounds of the speaker. The limitations of alternate means of reducing noise relative to speech were discussed. A bone-conductance microphone was proposed and its audio characteristics measured and compared to conventional, good quality microphones. Its principal advantages are shown to be high sensitivity to sounds generated by the *S*, and a superior ability to exclude noise from a signal.

APPARATUS NOTE

THE GRAYSON-STADLER ELECTRONIC SWITCH IN REACTION-TIME MEASUREMENTS

A review of several studies which have used the Grayson-Stadler electronic switches (Series 829) in reaction time (*RT*) measurement has indicated that errors exist in those studies. This note calls attention to the characteristics of these switches which may result in errors in the measurement. These characteristics are described in the instruction manual for these switches under the section dealing with "Periods," but their relevance for *RT*-measurement may not be immediately apparent.

Whenever the Series 829 switches are given the command to switch a signal from one channel to another, there is a delay between the command to switch and the beginning of the switching action. This delay is approximately equal to the rise-decay time set on the front panel of the switch. This delay occurs regardless of the mode of triggering the switch, *i.e.* either internally or externally. In *RT*-studies, the switch is usually externally triggered in the following manner. A set of relays, frequently contained in or controlled by an interval-timer, triggers the switch by making a momentary contact between one of the external control pins and the ground. These relays simultaneously start the *RT*-clock.

With such simultaneous triggering of the switch and clock, the intent is to measure the *RT* to the signal as it is switched-on. As indicated above, however, if there is a rise-decay time set on the switch, the actual switching of the signal will be delayed by a duration approximately equal to the rise-decay time. For instance, if the rise-decay time is set equal to 250 m.sec., there will be a 250 m.sec. delay between the command to switch the signal on, and the beginning of the switching action. Under such circumstances, the *RT*-clock will count 250 m.sec. before the beginning of the signal presented to *S*, and *S*'s *RT* will equal his actual *RT* to the slow-rise signal, ***plus*** the 250 m.sec. delay-time between the command to switch and the initiation of the switching action! A similar delay occurs at the command to switch the signal off, and hence if *RT*s are measured contingent upon the offset of the signal, and the switch and *RT*-clock are simultaneously triggered, offset *RT*s will be artifactually inflated by a magnitude equal to the decay time.

In a personal communication with Mr. Rufus L. Grayson, of the Grayson-Stadler Co., he suggests that this difficulty might be overcome by introducing an external timer which could be adjusted to be equal to the delay (itself approximately equal to the rise-decay time) and using this timer to trigger the *RT*-clock at the appropriate time following the external triggering of the switch. A second alternative would be to use an external Schmidt trigger circuit driven from the output of the electronic switch. The "threshold" of the Schmidt trigger could be set at some point on the rise (or decay) slope of the signal being switched. Finally, and perhaps most simply, the switching delay might be determined by means of an oscilloscope and this constant could then be subtracted from the *RT*-data for each rise-decay time used.

University of Pittsburgh
THOMAS G. STICHT

NOTES AND DISCUSSIONS

A COMMON PRINCIPLE RELATING TO THE MÜLLER-LYER AND PONZO ILLUSIONS

Müller-Lyer illusion. The Müller-Lyer illusion is a figure in which the apparent-length of a straight line is distorted when arrow-heads are added in such a way as to bracket and bisect it.[1] This illusion is illustrated in Fig. 1. It will be seen clearly that the part of the horizontal line bounded by ingoing arrow-heads appears to be much shorter than that bounded by outgoing arrow-heads. The illusion persists when modified to the form of two individual lines one bracketed by ingoing, the other by outgoing arrow-heads. Research concerned with features of this particular figure probably exceeds the volume of that undertaken upon all the other geometrical illusions taken together.

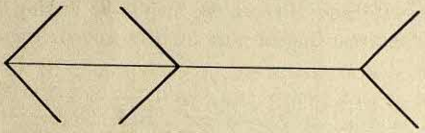


FIG. 1. THE MÜLLER-LYER ILLUSION

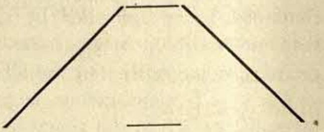


FIG. 2. THE PONZO ILLUSION

The Ponzo illusion. The illusion in which two lines, each of equal length, placed one above the other within two lines inclined towards each other, undergo distortion in such a way that the lower line appears to be shorter than the upper line, was first investigated by Ponzo.[2] This illusion is illustrated in Fig. 2. It may be described as an apparent-distortion of magnitude induced by the presence of the inclined lines.

Many suggestions have been made in the attempt to relate the apparent distortions seen in these figures to a common principle. No theory proposed hitherto has been successful, however, in giving a consistent and comprehensive account of these unusual and diverting phenomena. The theme common to many attempted explanations of perceptual distortions, from Thiéry to Gregory, is that which implies them to arise as a consequence of the tendency of abstract, two-dimensional patterns to be interpreted as concrete, three-dimensional scenes.[3] A difficulty encountered by all theories which understand illusory figures to be extended in apparent depth, however, is that they imply that each half of the shaft of the Müller-Lyer figure should be seen at different apparent depths at one and the same time. While this assumption may be acceptable when considering each component part of the figure seen independently, it is less convincing when extended to consideration of the composite figure. The reason for this is that the distance of the reference plane set by the ends of the bisecting arrow-head must be seen as being simultaneously both nearer to, and farther from, the observer than the shaft if a consistent three-dimensional interpretation of the figure is to be maintained. Difficulties similar to this also arise when attempting to understand the distortions seen in other illusory figures in this way.

---

* The research reported in this paper was supported by Research Grant No. B/H/139 from the Science Research Council, and No. 70/GEN/9617 from the Ministry of Defence.

[1] F. C. Müller-Lyer, Optische Urteilstäuschungen, *Arch. f. Physiol.*, Suppl. Bd., 1889, 263-270.

[2] Mario Ponzo, Rapports entre quelque illusions visuelles de contraste angulaire et l'appréciation de grandeur des astres à l'horizon, *Arch. Ital. d. Biol.*, 58, 1913, 327-329.

Two Ponzo figures juxtaposed. The Ponzo illusion appears to bear a superficial resemblance to the Müller-Lyer figure since each embodies a pattern of inclined lines within its configuration. The similarity between these two illusions was noted by Brentano but he was unable to develop a principle according to which the distortions seen in each could be explained satisfactorily.[4] The distortion predicted by the Ponzo principle appears, at first sight, to be restricted to lines placed in orientations orthogonal with the shaft of the Müller-Lyer figure rather than to the shaft itself. It is possible, however, that preoccupation with the influence of the arrow-heads as such has served to obscure a hitherto unnoticed feature of their influence upon the distortion which appears in this figure. This may be demonstrated by placing one Ponzo illusion alongside another in such a way that the lower end of the right-hand oblique in the left figure meets the upper end of the left-hand oblique in the right figure. The two illusions juxtaposed in this way are illustrated in Fig. 3. In each side of this figure the spatial distortion characteristic of the Ponzo illusion appears clearly, the upper member of each of the two pairs of horizontal lines being apparently longer than the lower member.

Four Ponzo illusions juxtaposed. The addition of two further oblique lines to Fig. 3 and of additional horizontal lines, one on either side, yields the pattern shown in Fig. 4. This drawing is defined by four Ponzo illusions, two as described above, and two further figures otherwise similar but inverted in orientation. In this figure the central line bounded by ingoing arrow-heads appears shorter than the outer lines, while that bounded by outgoing arrows appears longer than the outer lines. Accordingly, the left-hand central line appears to be foreshortened in relation to that on the right. It will be appreciated readily that the pattern illustrated in Fig. 4 is a special case of the Müller-Lyer illusion. Clearly, the nature of the distortion in this figure is in accord with that seen in the Ponzo illusion. Hence, it seems justifiable to conclude that the spatial distortions appearing in both these figures are related by common principle.

---

[3] Armand Thiéry, Über geometrisch-optische Täuschungen, *Phil. Stud.*, 12, 1896, 67-125; R. L. Gregory, Distortion of visual space as inappropriate constancy scaling, *Nature*, 199, 1963, 678-680.

[4] Franz Brentano, Über ein optisches Paradoxen, *Z. Psychol.*, 3, 1892, 349.

Ponzo illusions with different spacings. A major source of concern for those who have conducted systematic studies of the magnitude of distortion

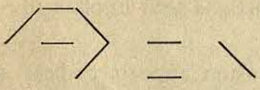


FIG. 3. TWO PONZO FIGURES JUXTAPOSED

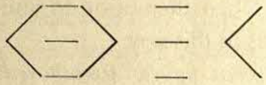


FIG. 4. FOUR PONZO ILLUSIONS JUXTAPOSED

evident in the Müller-Lyer figure has arisen from the fact that ingoing and outgoing arrow-heads subtending the same angle appear to induce distortions which differ in extent. Naturally, the distortions are in opposite directions, one apparently foreshortening the shaft, the other elongating it, but since the angles subtended by the brackets are the same there seems no obvious reason why the absolute extent of each influence should differ. This difference was first noted by Binet and has been reaffirmed by Pollack, by Pollack and Chaplin, and by Lucas working in the author's laboratory.[5] It is possible that failure to explain the discrepancy between the extents of the illusion in the two directions again derives from a predilection to consider the angle of the arrow-heads as such rather than the relevance of the lateral relationship between the oblique lines. As is described above, the distortion seen in the Müller-Lyer figure may be predicted from that which characterises the Ponzo illusion. While the angles of the arrow-heads bracketing and bisecting the shaft of the figure are the same, however, the separations between the two sets of oblique lines are different, in that bounded by outgoing arrow-heads. This difference in separation is illustrated in Fig. 5(A) showing the separation of the oblique lines in the being closer in the part of the figure bounded by ingoing arrow-heads than part of the figure usually associated with acute brackets and (B) that in the part associated with obtuse brackets.

---

[5] Alfred Binet, La mesure des illusions visuelles chez les enfants, *Rev. Phil.,* 40, 1895, 11-25; R. H. Pollack, Simultaneous and successive presentation of elements of the Müller-Lyer figure and chronological age, *Percep. mot. Skills,* 19, 1964, 303-310; R. H. Pollack and M. R. Chaplin, Effects of prolonged stimulation by components of the Müller-Lyer figure upon the magnitude of illusion, *Percep. mot. Skills,* 18, 1964, 377-382; Ann Lucas, *An Experimental Study of the Geometrical Illusions,* 1967, Doctoral dissertation, University of Newcastle upon Tyne, England.

Ponzo illusions superimposed. In Fig. 6, the two Ponzo illusions shown in Fig. 5 (A) and (B) have been superimposed in such a way as to define a larger figure with three rather than two enclosed lines. This figure suggests that the figure usually described as the Ponzo illusion is a special case of a more general distortion in which a gradient of attenuation appears within areas enclosed by inclined lines. The nature of this distortion-gradient is such that horizontal lines more closely adjacent to the locus of intersection of inclined lines appear to be increasingly longer than those which are situated further from it. Application of this principle to a wide sample of figures, commonly described as 'the angle illusions,' suggests that a general explanation of the distortions seen in them may be developed in these terms. As far as we are aware, however, the function describing the magnitudes of distortion-gradients in figures of this kind remains to be evaluated and it is intended that this shall provide the subject of a further paper.

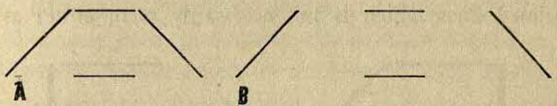


FIG. 5. PONZO ILLUSIONS WITH DIFFERENT SPACINGS BETWEEN THE OBLIQUE LINES

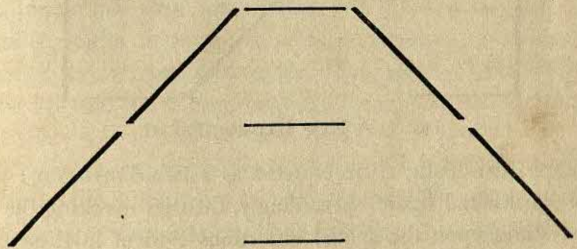


FIG. 6. THE PONZO ILLUSION SUPERIMPOSED

It should be noted that, while relating these two illusions in principle, the considerations outlined above fail to overcome the objection raised by any incompatible depth features which may become associated with the ends of the central-fin of the Müller-Lyer figure. If appropriate, the distortion predicted by the distortion-gradient principle should be relevant to a large number of other spatial patterns, however, one example of which is illustrated in Fig. 7. Here the outer bracketing lines are drawn in a vertical orientation, the central line only being inclined. The four horizontal lines are again all equal in length. The illusion appears in this figure in the direction predicted, the upper-left and lower-right horizontal lines being apparently longer than the lower-left and upper-right horizontals respectively. Further, any depth features in this figure are minimal in comparison with the Müller-Lyer and Ponzo illusions. Hence, it seems reasonable to suggest that spatial distortions of this kind are truly 'geometrical' in the sense that they appear within any appropriate two-dimensional figure containing inclined lines which is not necessarily seen, either as being extended in apparent-depth, or interpreted as a two-dimensional projection of a three-dimensional figure. Accordingly, theories invoking the influence of depth features upon the spatial distortions evident in these figures do not appear to be entirely convincing.

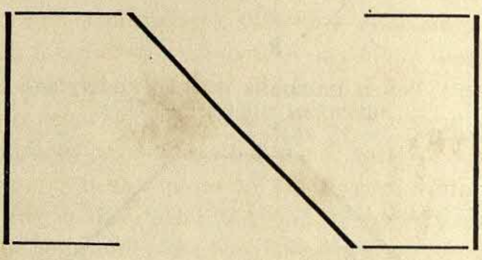


FIG. 7. A NEW ILLUSION-FIGURE

SUMMARY

The principle underlying the spatial distortion in the Ponzo illusion is applied to the Müller-Lyer figure. It is suggested that the distortion seen in the Müller-Lyer figure is predictable from that evident in the Ponzo illusion if the adjacent oblique lines are considered as being inclined towards, or away from, each other rather than as converging to the form of arrow-heads. Understood in this way, the differential extent of distortions apparent in lines bounded by arrow-heads subtending the same angle, but flaring in opposite directions, appears to be explained. The further suggestion is made that the Ponzo illusion is a special case of a distortion-gradient which appears in two-dimensional, rectilinear figures embodying patterns of inclined lines. The nature of this distortion-gradient is such as to

imply that the geometrical illusions are truly 'geometrical' since it becomes evident in any appropriate figure which is not necessarily interpreted as having extension in depth.

University of Newcastle upon Tyne — GERALD H. FISHER
England

CONCEPTUAL COMPLEXITY, RECALL, AND IDENTIFICATION

Conceptual complexity, as a dimension of personality, has been shown to account for variance in emerging group-structure, in subjects' perceptions of their environmental tasks, in informational search, and in attitudinal change.[1] The Paragraph-Completion Inventory,[2] a measure of conceptual complexity frequently used, presumably differentiates individuals in terms of their ability to *interrelate* aspects of a reasonably complex environment. One important claim implicit in Schroder *et al.*'s theory is that this greater "integrative" capacity of conceptually complex individuals does *not* arise from a possible superiority in the availability of informational units to them. This contention—that conceptually simple and conceptually complex *S*s, if found to differ in measures of informational processing, will not differ in measures of mere storing (*i.e.* their recall of the informational units involved)—has never been subjected to a direct test. One purpose of the present study was to provide such a test.

A second purpose was to present further evidence for the validity of the conceptual complexity-construct by investigating its predictiveness in a newly designed environmental task and by using more objective measures of performance than those used in previous researches involving the "War Game."[3] The performance measure used in this study is one of identification. *S* is required to identify an object inferentially through using various dimensional values in combination.

Method: (1) Subjects. Twenty-four *S*s, 12 conceptually simple and 12 conceptually complex, were used in this study. Conceptual complexity was determined by the Paragraph Completion Inventory.[4] The two personality groups were matched in age

---

* This research was supported by the Office of Naval Research under Contract No. 1858(42). We thank Dr. Bertram L. Koslin for helpful suggestions.

[1] H. M. Schroder, M. J. Driver, and Siegfried Streufert, *Human Information Processing: Individuals and Groups Functioning in Complex Social Situations*, 1967, 101; 114; 144.

[2] Schroder *et al., op. cit.*, Ch. 12.

[3] Siegfried Streufert and H. M. Schroder, Conceptual structure, environmental complexity, and task performance, *J. exp. Res. Person.*, 1, 1965, 132-137.

[4] Schroder *et al., op. cit.*, Ch. 12.

(ranging from 16 to 34 yr.), and on intelligence, using the Wonderlic Personnel Test which correlates 0.87 with the Otis Test of Intelligence.[5]

(2) Procedure. Written instructions gave *S* a first orientation to the task—the "Anti-Air Warfare"—and to the part he was to play in it: He was to assume the role of commander over a fleet moving slowly toward enemy-held territory. Activity would be possible at any moment. His objective was to detect the enemy while keeping his own fleet undetected. Information about the components of his imaginary task-force, including particularly the operating characteristics of his search, was printed on the back of IBM cards (101 in number) containing one or two sentences each. *S* was given 45 min. to familiarize himself with this task.[6]

As a *control,* which then followed, *S* was asked to take a timed recall-test consisting of 12 items, (*e.g.* "What is the speed of neutral aircraft?"). This task can be conceived of as a control with the purpose of making sure that any differences between the conceptually simple and the conceptually complex *S*s in the experimental tasks (see below) could not be accounted for by differences in their recall of the informational units to be processed. After completing the test, *S* was presented with a list of the right answers to the test-items.

The first experimental task consisted of a set of problems of identification. It took the form of a timed test containing 12 items.

Instructions. The instructions were as follows: You will now be given a few problems in which your task is to infer the type of object from a given set of characteristics. For each of the problems, please indicate which of the six possible foreign objects is most likely the one that has the given characteristics. Note that in these problems it is assumed that you are given complete information concerning the characteristics of the objects (*e.g.* speed, altitude). In reality, such complete information may not always be available due to the limitations of your search-apparatus. [For example, the first item read: "One object; at a speed of 250 m.p.h.; 9,000 ft. high; in a straight flying pattern; with active radar."]

S was allowed to refer to the information-cards which he had received previously. After 5 min. he was asked to stop and read through a sheet containing the right answers to the problems.

The *second task* consisted of a "detection-exercise." This task was a simulation of an ongoing tactical situation. *S*'s specific task in this exercise was to identify the source of a given radar signal, that is, to identify each foreign object whose presence was indicated on a radar map (a simulation of a radar screen). Twelve such radar maps constituted the total experimental inputs. They were presented over a period of 48 min. at regular intervals. Each map contained a (mimeographed) geographical representation of the area of operations. Superimposed on it were the radar signals —a map could contain from three to seven signals—thus giving each map the function of a situational report covering the preceding time-interval. In evaluating a given situational report, *S* was allowed to refer to previous reports. Thus one important difference between the first and the second experimental task was that in the Detection Exercise, *S*'s information-processing required him to *interrelate* the various informational inputs he had received.

---

[5] Published by E. F. Wonderlic, P. O. Box 7, Northfield, Illinois.

[6] A revised version of this task-environment has been described in detail in Helmut Lamm, The effect of conceptual complexity and group discussion on risk taking and judgmental behavior. Unpublished Doctoral dissertation, Princeton University, 1966.

The performance measure used for the analysis was the number of signals, over all inputs (situational reports) on which *S* failed accurately to identify the type of object.

Hypotheses. In the light of the procedure described above, our hypotheses may be summarized as follows. There will be no difference between conceptually Simple and Complex *S*s on the recall-task; by contrast, the *S*s high in conceptual complexity will be superior on the identification-test as well as on the detection-exercise.

Results. As Table I shows, *t*-tests between the two personality groups showed no differences on the recall-test, but significant differences on the two experimental tasks ($t = 1.72$, $p < 0.05$ for the Identification-Problems; and $t = 1.86$, $p < 0.05$ for the Detection-Exercise). As predicted, the conceptually complex *S*s were superior on each of the two information-processing tasks.

TABLE I

COMPARISONS OF CONCEPTUALLY SIMPLE AND COMPLEX *S*S ON THE RECALL-TASK AND ON THE TWO INFORMATION-PROCESSING TASKS

	Conceptual complexity					
	High (N=12)		Low (N=12)			
	X	*SD*	*X*	*SD*	*t*	*p*
Recall-task*	5.33	2.21	6.42	3.55	0.86	n.s.
Identification-task†	10.00	2.16	8.42	2.18	1.72	0.05§
Detection-exercise‡	20.92	9.56	27.83	7.81	1.86	0.05§

* Units of information accurately recalled.
† Number of identifications accurately inferred from list of characteristics.
‡ Number of instances *S* failed to identify the objects from the 'radar' signals.
§ One-tailed.

Discussion. Recent reviews of the literature on problem-solving have pointed to the need for further inquiry into the role of personality variables in problem-solving.[7] The findings of the present study clearly indicate that conceptual complexity as a personality variable accounts for differences in information-processing that contains features of inference-making (one aspect of directed thinking). Intelligence could not account for these differences. Nor can memory factors account for the present findings, since no differences in the recall of relevant informational items were found. In view of the current tendency of investigators to approach the study of problem-solving in terms of information-processing, it may be concluded that the information approach to personality, characterizing the theory of Schroder and his associates may indeed prove fruitful for studying the interaction of situational and dispositional variables in problem-solving. This conclusion is supported by the present findings. Further studies along these lines should inquire into the extent to which conceptual complexity accounts for variance in a variety of other tasks lending themselves to objective performance-measures. Shaw's proposed taxonomy of tasks is but one of recent efforts at systematic ordering of more widely used tasks which could be utilized in relevant research.[8]

---

[7] *E.g.* R. M. Gagné, Problem solving, in A. W. Melton (ed.), *Categories of Human Learning,* 1964, 294-317.

Summary. *S*s operating at complex and simple levels of conceptual complexity were familiarized with the technical problem of a Task Force, and then each was asked to play the role of a navy commander, requiring him to identify unknown foreign objects on the basis of clues presented at regular intervals. It was found that conceptually complex *S*s did not differ from simple *S*s in their recall of informational units relevant to the task, but were able to make more accurate identifications. The findings are taken as further support for the basic tenets of conceptual complexity theory and are in particular seen as a confirmation of a hitherto untested implication of that theory.

Princeton University

HELMUT LAMM
THOMAS L. COFFMAN
MARVIN KARLINS
HAROLD M. SCHRODER

HETEROCHROMATIC STABILIZED IMAGES: A CLASSROOM DEMONSTRATION

When the image of a colored disk surrounded by a ring of a different color is stabilized on the retina, the disk soon disappears; that is, *O* reports that he sees a single large disk whose color is that of the ring.[1] The phenomenon is easier to deal with conceptually, and perhaps experimentally, if the outer border of the ring is unstabilized (for example, by placing a suitable stop in the optical system), but it can be obtained as well if the ring is surrounded by darkness. The effect is perceptually similar to the "filling in" that occurs when objects are imaged across the blind spot, and perhaps the ultimate explanations of both phenomena will prove to be similar. Actually, however, the stabilized-image effect is the more challenging theoretically, because the 'filling in' occurs over a retinal region, in which there is no gap in the receptor-surface; it occurs, in fact, in the region of greatest acuity. The purpose of the present note is to describe how the phenomenon may be observed without the elaborate apparatus used in stabilized-image experiments.

---

[8] M. W. Shaw, *Ann. Tech. Rept.*, University of Florida, Gainesville, Florida, 1962.

[1] A. L. Yarbus, The perception of an image fixed with respect to the retina, *Biophysics,* 2, 1957, 683-690; John Krauskopf, Effect of retinal image stabilization on the appearance of heterochromatic targets, *J. opt. Soc. Am.*, 53, 1963, 741-744; H. J. M. Gerrits, B. deHaan, and A. J. H. Vendrik, Experiments with retinal stabilized images. Relations between the observations and neural data, *Vis. Res.*, 6, 1966, 427-440.

A certain mystique has developed around stabilized images, with the implication that perception under stabilized conditions is somehow qualitatively distinct from normal perception. There really is no evidence for this view. Under ordinary viewing conditions, some receptors receive temporally varying inputs as eye-movements shift the position of contours (or rather, the gradients in the images of contours) on the retina. Maintained vision of objects seems to depend on this variation in retinal stimulation, which can be reduced in several ways, stabilization being only one. Another effective procedure is to decrease the luminance-gradients of contours. Cornsweet has used this method effectively to demonstrate the disappearance of achromatic displays to large audiences.[2] A white disk is projected on a screen and then the projector is thrown out of focus. The viewers are told to fixate a small fleck on the screen. Most members of the audience soon report that the previously visible disk has disappeared. The methods that may be used for heterochromatic targets are similar to Cornsweet's, but there are some further details to be considered.

The most satisfactory procedure found to date requires a stereo-projector which permits good control over the registry of two separate slides. Alternatively, the slides may be put into two projectors side-by-side. One slide shows a clear disk with an opaque surround; the other, an opaque disk with a clear surround—the negative of the first slide. A special projection screen is arranged by placing a black screen with a circular hole in its center in front of an ordinary screen. The black part of the screen serves to produce a steep gradient in luminance at the outer border of the ring, simulating the effect of the destabilizing stop in the stabilized-image apparatus. Any heterochromatic disk-annulus combination desired can be obtained by placing suitable color-filters in front of the two projection-lenses. No special care is required in selecting the filters, but optimal disappearance-effects are obtained when the luminances of the disk and the annulus are equated. Control of luminance with the aid of neutral filters is, therefore, helpful.

---

[2] T. N. Cornsweet, Stabilized image techniques, in M. A. Whitcomb (ed.), *Recent Developments in Vision Research,* National Academy of Sciences, 1966, 171-184.

To demonstrate the disappearance of the disk, the ring and disk are first sharply focused on the screen, centered in the clear-screen area, and brought into good registry. Both projectors then are defocused. The *O*s are told to fixate steadily on the center of the display; although a fixation-point is not necessary, a small black dot in the center of the screen is helpful. Most *O*s then report the disappearance of the disk. If the demonstration is performed in a long room, those in the front will be more likely to observe the disappearance, since the retinal gradients will be shallower for them than for others farther from the screen.

A second way to demonstrate the phenomenon requires only one projector. A single in-focus slide is prepared of the desired heterochromatic pattern. This slide is projected out of focus, on the compound screen described above. Since no differential control over the luminance of the disk and annulus is possible after the slides are prepared, the equation must be taken care of during the preparation of the slides.

A third method, also requiring but a single projector, is to take various out-of-focus pictures of a heterochromatic disk-annulus pattern. They may be mounted together with an opaque circular hole, thus making a special screen unnecessary. Again, luminance-adjustments must be made when the pictures are taken. One difficulty with this procedure is that, apparently because of the nonlinearity of the photographic process, a ring appears at the border between the disk and the annulus. The color of the ring is that which is expected from color-mixture theory. Nevertheless, it is possible with this procedure to observe the same disappearance as with the other procedures.

A complete explanation of the phenomenon cannot be given. Roughly speaking it seems as though information concerning discontinuities between regions determines how the regions themselves are seen. Under ordinary circumstances, the playing of contours over the receptors provides information about the discontinuities. When this variation in stimulation is attenuated by stabilization or blurring, the information about certain contours is lost, and the remaining information determines the visual outcome. The only information remaining concerns the transition from the black surround to the annulus. Thus, the whole target is seen in the color of the annulus.

Bell Telephone Laboratories JOHN KRAUSKOPF

TILT AND CENTRIFUGATION IN CHANGING THE DIRECTION OF BODY-FORCE

An increasing interest in the determinants of spatial orientation has resulted in numerous studies concerned with postural and visual orientation during a change in the direction of force acting on the body.[1] The direction of force can be changed either by tilting the body relative to a fixed direction (*e.g.* gravity) or by rotating *O* in a centrifuge. In the latter case the direction of the force is the resultant of the gravitational and centrifugal components and is referred to as the gravito-inertial vertical. In short, the direction of the force can be varied by tilting the body with the direction of force constant or by modifying its direction with body-orientation constant.

An observer (*O*) cannot distinguish between the effects of these linear accelerations.[2] Howard and Templeton have argued that these two procedures are identical in that their effects on any object or person are the same in terms of force-direction. Thus they considered that centrifugation is an irrational and inefficient method when body-tilt has the same effect. Defence of the use of the human centrifuge has concerned only cognitive aspects associated with the procedure.[3]

The physics of the two operations is clear. When considering only the direction of force, it is identical. The psychophysics, however, is not so straightforward. Tilting the body relative to gravity generally results in the E-phenomenon for angles less than about 60° and the A-phenomenon for larger angles.[4] Changing the direction of force by centrifugation results in the A-phenomenon for angles less than about 20° and the E-phenomenon for larger angles, when the effects are measured from the gravito-inertial vertical. Further, the magnitude of the effects under centrifugation are smaller than under body-tilt and there is a marked lag-effect under the former condition not found with the latter.[5]

---

[1] E. F. Miller, II, and Ashton Graybiel, Role of the otolith organs in the perception of horizontality, this JOURNAL, 79, 1966, 24-37; Magnitude of gravitoinertial force, an independent variable in egocentric visual localization of the horizontal, *J. exp. Psychol.*, 71, 1966, 452-460.

[2] L. B. W. Jongkees and J. J. Groen, The nature of the vestibular stimulus, *J. Laryng.*, 61, 1946, 529-541; I. P. Howard and W. B. Templeton, A critical note on the use of the human centrifuge, this JOURNAL, 76, 1963, 150-152; *Human Spatial Orientation*, 1966, 197-200.

[3] H. A. Witkin, Uses of the human centrifuge in studies of orientation of space, this JOURNAL, 77, 1964, 499-502.

[4] Hermann Aubert, Eine scheinbare bedeutende Drehung von Objecten bei Neigung des Kopfes nach rechts oder links, *Arch. path. Anat. Physiol.*, 20, 1861,

The otolithic organs are involved in both these effects.[6] These organs respond with a shearing action to that component of the force which acts laterally. The effective stimulus for mechanical shearing of the otoliths is a force acting tangentially to the macular surface.[7] Centrifugation increases the magnitude of force acting on the body as well as changing its direction. Resolution of the force for tilt and centrifugation yields a greater lateral component (L) for the latter condition when the direction of force acting on the body is identical, as shown in Fig. 1. Thus although

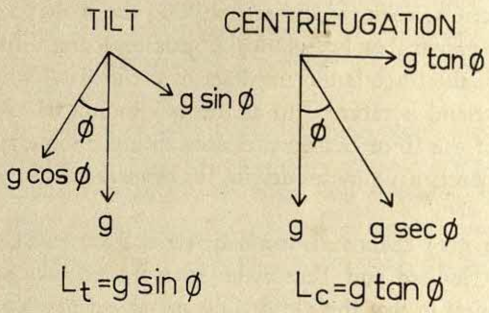


FIG. 1. RELATIVE MAGNITUDE OF FORCE COMPONENTS ACTING AT RIGHT ANGLES TO BODY AXIS FOR TILT (L_t) AND CENTIFUGATION (L_c). The directions of resultant force are identical for the two conditions.

381-393; Martin Bauermeister, Effect of body tilt on apparent verticality, apparent body position, and their relation, *J. exp. Psychol.*, 67, 1964, 142-147; E. F. Miller, II, A. R. Fegly, G. van den Brink, and Ashton Graybiel, Egocentric visual localization of the horizontal as a function of lateral body tilt up to 90° with respect to gravity, *Bu. Med. Proj. MR005.*, 13-6001, Naval School of Aviation Medicine, August 1965; G. E. Müller, Ueber das Aubertsche Phänomen, *Z. Sinnesphysiol.*, 49, 1916, 109-244.

[5] Ashton Graybiel and Brant Clark, Validity of the oculogravic illusion as a specific indicator of otolith function, *Aerospace Med.*, 36, 1965, 1173-1181.

[6] Brant Clark and Ashton Graybiel, Perception of the visual horizontal in normal and labyrinthine defective observers during prolonged rotation, this JOURNAL, 79, 1966, 608-612; Miller and Graybiel, *op. cit.*, 1966, 24-37; Hermann Schöne, On the role of gravity in human spatial orientation, *Aerospace Med.*, 35, 1964, 764-772.

[7] D. E. W. Trincker, The transformation of mechanical stimulus into nervous excitation by the labyrinthine receptors, in *Biological Receptor Mechanisms, Symp. Soc. exp. Biol.*, No. 16, 1962, 289-314.

the physics of the two operations yields an identity in terms of the direction of force acting relative to the body, the lateral components are not identical, and it is these that are physiologically significant to the functioning of the otolithic organs. For small angles (less than 20°) the lateral forces are approximately equal. With body-tilt the maximal lateral component for the utricular otoliths is reached when *O* is recumbent ($\phi = 90°$, $L_t = g$). Under centrifugation the same lateral force occurs with the gravito-intertial vertical at half this value ($\phi = 45°$, $L_C = g$). For angles greater than 20° the tangential values become increasingly greater than the sinusoidal, and therefore the lateral force becomes increasingly greater. Considered in relation to the force acting on the otolithic organs the two conditions become increasingly discrepant. This point was appreciated by Woellner and Graybiel. Measuring ocular countertorsion under conditions of body tilt and centrifugation yielded separate functions when plotted in terms of the direction of force, but the same function when compared in terms of lateral force.[8] Furthermore, maintaining the same angle of body-tilt relative to increasing gravito-inertial forces increases the magnitude of the E-phenomenon.[9]

Physical identity of the effects of body-tilt and centrifugation with respect to the direction of force has been employed as an argument against the use of the human centrifuge.[10] It is here argued that this identity obscures the fact that the lateral component of the resultant force (whether gravitational or gravito-inertial) is physiologically of greater importance and is not identical under the two conditions. The otolithic receptors of the utricle (and saccule) respond to this lateral force. The human centrifuge is, therefore, a legitimate method of varying this component of the force in a manner otherwise impossible. It is not an irrational and inefficient method of studying spatial orientation under modified directions of force.

Monash University
Melbourne, Australia

N. J. WADE
R. H. DAY

---

[8] R. C. Woellner and Ashton Graybiel, Counterrolling of the eyes and its dependence on the magnitude of gravitational or inertial force acting laterally on the body, *J. Appl. Physiol.*, 14, 1959, 632-634.

[9] Miller and Graybiel, *op. cit.*, 1966, 452-460.

[10] Howard and Templeton, *op. cit.*, 1963, 150-152.

THE SIZE-DISTANCE INVARIANCE HYPOTHESIS: A REPLY TO COOK

Cook considered previous tests of the size-distance invariance hypothesis inexact because they relied on the doubtful assumption that the apparent sizes and distances of the comparison objects were identical with their physical sizes and distances.[1] While errors resulting from this assumption were probably small, at least when the objects of comparison were presented under full cue-conditions, Cook thought it necessary to devise a new technique which dispensed with this assumption. Cook's method, however, was based on two other implicit assumptions of questionable validity which, if false, would invalidate the method and could cause even greater experimental error.

Firstly, it was assumed that the standard and each of the comparison-objects appeared to lie entirely in the frontal plane on all viewing occasions and never appeared to be inclined at an angle to the frontal place or distorted in depth, for example, curved, bent, or corrugated. If an *O* saw the standard or one of the objects of comparison as inclined or distorted in depth, the apparent length of the object should have been longer than if it appeared completely frontal at the same distance, so that Cook's equations would not hold in this case. Although direct evidence bearing on the validity of this assumption, under Cook's experimental conditions, is unavailable, the writer has found a small number of *O*s who report illusory distortions in depth under similar conditions. Flat objects physically tilted backwards appeared curved to some *O*s and appeared slanted about a vertical axis to others, when the objects were viewed monocularly against homogeneous surrounds. Similar distortions may have occurred in Cook's experiment invalidating the first implicit assumption and rendering the test inaccurate. A post-experimental interview or questionary could be used to isolate and reject any *O*s who did not see each stimulus-object entirely in the frontal plane on all viewing occasions, and so ensure that the first assumption was not violated.

Secondly, the appearance of both the standard and each of the comparison-objects was assumed unchanged when viewed at different times by the same *O* if the physical values of the objects were not altered. Thus, the apparent size and also the apparent distance of the standard were assumed to be the same at the time *O* judged its size as when judgments of its distance were in progress. Also, the apparent size of the distance variable, during the period it was used to judge distance, was assumed to be identical with its apparent size when compared with the size-variable. But both the standard and distance-variable were viewed monocularly against homogeneous surrounds and it is possible that in the absence of reliable depth cues, the appearance of one or both of these objects changed from the first occasion to the second, due to the different types of judgment made, size at one time and distance on another. Since the validity of the second assumptions is in doubt, confidence cannot be placed in the accuracy of Cook's results.

---

[1] M. L. Cook, An exact test of the size-distance invariance hypotheses, this JOURNAL, 79, 1966, 568-575.

The problem of possible change in the appearance of the standard and variables could be overcome by using an indicator which forced *O* to judge size and distance simultaneously. For example, a number of comparison objects could be presented which varied in both size and distance and *O* could select the one which was most like the standard in size and distance. (Beck and Gibson used this type of indicator when testing the shape-slant invariance hypothesis to ensure that each pair of shape and slant responses were based on the same percept of the standard.[2] A testable equation relating the physical size and distance of the standard to the size and distance response can be derived for this type of indicator without assuming that the apparent properties or the standard of comparison-objects are identical with their physical properties, or that these objects are seen consistently from time to time, as is shown below.

Let s_{st} be the size of the standard, d_{st} its distance from *O* and s_{st}' and d_{st}' it apparent size and distance at any one time. Then from Cook's Equation [2],

$$s_{st}/d_{st} = s_{st}'/d_{st}' \qquad [1]$$

Now if s_{va} and d_{va} are the size and distance of the variable and s_{va}' and d_{va}' are the apparent size and distance of the variable at any instant, it follows by the same argument that

$$s_{va}/d_{va} = s_{va}'/d_{va}' \qquad [2]$$

At the moment that *O* makes his match,

$$s_{st}' = s_{va}' \qquad [3]$$

and

$$d_{st}' = d_{va}' \qquad [4]$$

---

[2] Jacob Beck and J. J. Gibson, The relation of apparent shape to apparent slant in the perception of objects, *J. exp. Psychol.*, 50, 1955, 125-133.

From Equations [1], [2], [3] and [4],

$$s_{st}/d_{st} = s_{va}/d_{va} \qquad [5]$$

Note that while the apparent size and distance values cancel to give the conventional expression for the size-distance invariance hypothesis in Equation [5], this equation is valid only under special circumstances, for example when a simultaneous match for size and distance is made.

Summary. Cook's technique eliminated one questionable assumption on which previous tests of the size-distance invariance hypothesis relied but incorporated two other doubtful assumptions which make the validity of his experimental test questionable. Alternative suggestions have been offered to eliminate the assumption that Cook avoided and to cope with the two doubtful assumptions on which his method is based.

The University of Western Australia ROGER BRITTEN

AVOIDANCE-LEARNING IN THE TIGER SALAMANDER

In view of the difficulty encountered in demonstrating learning in amphibians, it is interesting that the tiger salamander (*Ambystoma tigrinum*), which has been studied anatomically for many years, rapidly learns a one-way avoidance-response—although not to electric shock. Whether the defined response simply is walking out of a chamber or climbing out of it, shock produces only erratic behavior. Hot air, which probably is aversive because of its evaporating effect, is a more adequate *US*, but has one fault: It first induces an approach-response, which gives way to an escape-response only later, as the animal becomes dry; the approach-component seems to retard acquisition of the avoidance-response, although all *S*s do eventually learn. A bright light does, however, reliably elicit low-latency escape-responses and many *S*s trained begin to avoid the light regularly within 15 trials.

The *CS-US* interval used in most of the pilot work was 7 sec., with 7-10 trials per day and an intertrial interval of 2 min. The apparatus includes a small negative chamber (4 in. long × 2 in. wide) leading to a larger positive chamber (6 in. square). The negative chamber is illuminated by a 100-w. lamp, the positive chamber is dark, and the two are separated by a swinging door. The (vibratory) *CS* is provided by a buzzer fixed to the substrate. The door is opened as the *CS* is initiated, and a white line between the two chambers provides a convenient criterion of response. Between trials the animal is swabbed gently with wet cotton.

Carleton University
Ottawa, Canada
A. JOSEPH RAY, JR.

FOURTEENTH ANNUAL MEETING OF THE SOUTHWESTERN PSYCHOLOGICAL ASSOCIATION

The fourteenth annual meeting of the Southwestern Psychological Association was held at the Hotel America, Houston, Texas, April 27-29, 1967. The program included 24 paper sessions, 15 symposia, graduate and undergraduate student competitions, a distinguished psychologist invited address by Arthur Melton, University of Michigan, on "Interference in short-term memory," and a Presidential Address by Malcolm Arnoult, Texas Christian University, on "Perception as behavior."

A special symposium honoring Distinguished Professor Karl M. Dallenbach was chaired by Wayne H. Holtzman, University of Texas. James V. McConnell, University of Michigan, presented a paper, "Learning in planarians," John Lacey, Fels Research Institute, presented, "Cardiovascular function and sensory-motor integration," and Harry Helson, Kansas State University, presented, "Cognitive factors in adaptation-level phenomena."

Lynda Uphouse, Austin College, won first prize in the special undergraduate competition for her paper, "Hunger, thirst, and sex as irrelevant drives upon each other." David Sisemore and Sabra S. Hassel, University of Arkansas, won first prize in the graduate competition for their paper, "The effect of variable and constant discriminative stimuli upon alternation behavior," and other prizes were won by Hugh V. Angle, Texas Christian University, for his paper, "The role of temporal variables controlling sequential interresponse time dependencies," and by John Schneider and Julian Burn, University of Oklahoma Medical Center, for their paper, "The relationship among skill-chance preferences, internal-external control, and signing-not-signing of attitude questionnaire." These competitions are sponsored jointly by SWPA and Psi Chi.

The Publishers Prize for Research Excellence for the best paper presented by a person whose Ph.D. was conferred after January 1, 1962 was won by Thomas L. Mentzer, Tulane University, for his paper "A species-specific factor in the stimulus control of behavior." The prize was presented this year by the Brooks-Cole Publishing Company.

New officers for the Association for 1967–68 are: President: Jackson B. Reid, University of Texas; President-Elect, Daniel Sheer, University of Houston; Secretary-Treasurer, Gladys Guy Brown, Dallas; Council Representative from Arkansas, Philip Trapp, University of Arkansas; Council Representative from Oklahoma, Oscar Parsons, University of Oklahoma Medical Center; with Edwin O. Timmons, Louisiana State University, and James Dyal, Texas Christian University, remaining as Council Representatives from their respective states.

Committee chairmen for the Association and for the convention were: Program, Dan Sheer, University of Houston; Membership, Paul L. Thomas, Austin College; Elections, Benjamin Fruchter, University of Texas; Local Arrangements, Marcus J. Fuhrer, Texas Institute for Research and Rehabilitation.

The 1968 meeting will be held at the Hotel Fontainebleau, New Orleans, Louisiana, April 18–20. Additional information may be obtained by writing Mrs. Betty Cleland at the Association's central office at 2104 Meadowbrook Drive, Austin, Texas.

University of Arkansas JOHN P. ANDERSON

Max Frederick Meyer: 1873—1967

Born June 15, 1873, Max F. Meyer died on March 14, 1967, at the age of ninety-four years. He was one of Carl Stumpf's first students at Berlin, where he took his Ph.D. in 1896. Like that of his teacher, his first and abiding interest was hearing and musical acoustics. He first published on the theory of hearing in 1899,[1] and then put forward his own theory in 1900, elaborating it in a monograph in 1907,[2] and again in a small book in 1950.[3] The lack of a suitable position in Germany or England caused him to seek to continue his work in America, but here again suitable posts in the more established universities of the east were not to be had. He accepted, instead, a position at the University of Missouri in 1900. He continued to write about the mechanism of the inner-ear in a form opposed to Helmholtz's resonance view, and also about theories of musical harmony and melody.

At the same time, he took his teaching duties and responsibilities as a Missouri faculty member very seriously and broadened his psychological concerns to almost all of experimental psychology and even abnormal psychology. In this period, for example, he wrote *The Psychology of the Other One,* an excellent statement of the behaviorist point of view, especially with respect to investigations with human beings.[4] It was accompanied later by *A Brief Manual of Psychology Demonstrations,* which was clearly intended as exposition to be used by the student of psychology.[5] These teaching books did not notably attract the attention of systematists, but yet they served American psychology well in formulating a behaviorism of observation and objectivity without resort to animals.

---

[1] M. F. Meyer, Zur Theorie des Hörens, *Arch. ges. Physiol.,* 78, 1900, 346-362.

[2] Meyer, An introduction to the mechanics of the inner ear, *Univ. Mo. Stud.,* 1907, 1-140.

[3] Meyer, *How We Hear; How Tones Make Music,* 1950, 1-117.

[4] Meyer, *The Psychology of the Other One,* 1921, 1-438.

He stayed at Missouri almost 30 years and his success as teacher and taskmaster has been heartily attested by such former students as O. H. Mowrer, P. R. Farnsworth, and A. P. Weiss. His career there ended with an unfortunate series of incidents attending a questionary, a series which in retrospect appears to have brought more shame on the Curators of Missouri University than on Meyer himself. He spent 1929 and 1930 in St. Louis at Central Institute for the Deaf continuing his research, and searching some applications for the education and testing of deaf children. He was then received by the University of Miami in Florida and the major part of his later life was spent in Miami, where his acoustic laboratory was in his home—sans grants, sans other subjects, and, to his delight, sans vacuum tubes.

Of course, Meyer's principal contribution concerned hearing and music, as he himself said. His model of the cochlear partition as a non-stretched membrane like chamois was never accepted in detail, but his opposition to the notion of stretched, tuned resonating sections was right. Some of the consequences of other aspects of his theory, which require, for example, the representation of frequency in the nervous system, probably cannot be supported in toto. His musical ear, coupled with his easy satisfaction that what he heard was incontrovertably fact, produced many articles, in polemic style that even in later years had to be softened by editors, this writer among them. He heard combination tones, combinations of tones, and Tartini pitches whose existence is still questioned by other authors. Presumably they should have been among his subjects.

American psychologists can rejoice that Meyer came here and must grieve at his passing. The enthusiasm and devotion to a basic idea, perhaps wrong in some aspects, is an admirable phenomenon that appears to be passing from the scene too rapidly. One has only to read some of the details, excellently put by Esper[6] to recognize a man of originality, persistent, and often misunderstood as an impolite crank in his adopted country, whose mark in the history of auditory theory and of behavioral theory is made and deeply grooved by decades of persuasion.

IRA J. HIRSH

Central Institute for the Deaf
St. Louis, Missouri

---

[5] Meyer, *A Brief Manual of Psychology Demonstrations,* 1922, 1-118.

[6] E. A. Esper, Max Mayer: The making of a scientific isolate, *J. Hist. Behav. Sci.,* 2, 1966, 341-356; Max Meyer in America, *ibid.,* 3, 1967, 107-131.

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

The Structure of Associations in Language and Thought. By JAMES DEESE. Baltimore, Md., The Johns Hopkins Press, 1965. Pp. xiii, 216. $6.50.

The aims of this book differ from those of two other major approaches to the study of free word associations. One line of research, popular in the early days of empirical study in this area, was concerned with a characterization of the relation between stimulus and response, in the hope (largely unfulfilled) that an ad hoc classification of the *Ss*' responses in relation to the stimuli that occasioned them would reveal the mental processes that gave rise to these responses. In contemporary psychology, the attempt to explain free association responses has been replaced by the use of free association data to explain other phenomena in the verbal learning laboratory.

Deese's goal is more ambitious than either of these two. He uses the free association technique to uncover psychologically significant relations in the (English) lexicon. The responses obtained in the test are not the focus of primary attention but, rather, serve merely as instruments for manifesting inherent relations among the stimuli. Such an objective, ably pursued by one of the leaders in the field, deserves the serious scrutiny of persons interested in the psychology of language and especially in its semantic aspects.

How can free association results shed light on the nature of semantic relations in the language? Deese's answer is essentially that this can be achieved through a proper analysis of the free association data. Giving a set of stimulus-words to a group of *Ss* and asking them to respond to each with the first word other than the stimulus that comes to their minds, one obtains a frequency-distribution of responses. Such distributions can be analyzed in a variety of ways, yielding ten or more different, but closely related, indices of word relatedness. Deese's measure, termed *index of commonality* (IC), bears a close affinity to these measures. For each pair of stimulus-words, it records the proportion of the number of actual common responses to the maximum number of possible common responses. In calculating the IC, it is assumed that a stimulus-word automatically elicits itself as a "representational response." Thus, although only a single response is obtained overtly, two responses are counted. As an example, consider the stimuli MOTH and INSECT. Of the 50 *Ss* who responded to these words (among others), 9 gave *fly* to both words, 2 gave *moth* to INSECT and 1 gave *insect* to MOTH. Counting the representational responses (50 *moth* to MOTH and 50 *insect* to INSECT), the number of common responses is 12. The maximum could have been 100, yielding a ratio of 0.12. IC values are calculated in this way between each stimulus and each other stimulus in a given set. IC matrices are then subjected to a factor analysis to extract common dimensions. Each factor identifies a group of words which belong together on the basis of their IC values. The loadings of every word on all factors fully describe the relations of this word with all other words in the matrix. The common variance of each word is thus broken down into factorial components.

The conception of words as composed of smaller entities is not new. It can be traced back at least to John Locke's 17th century analysis of complex ideas into their constituent simple ideas. Contemporary linguists of different theoretical persuasions and interests commonly analyze words into their components or features. For instance, *boy* in the feature analysis is described, in part, as a noun with the specification of *common, count, animate,* and *human;* and *father* in the componential analysis is characterized as *lineal kin, male,* and *one generation above ego.* Psychologists working with the semantic differential also describe words by placing them in reference to three general factorial axes: evaluation, activity, and potency. Deese's factorial analysis can be characterized in the following way in relation to these conceptions. (a) The present classification of each word is empirical, as in the case of the semantic differential, rather than logical or linguistic, as in the feature and componential analyses; (b) The classification is continuous rather than discrete, *i.e.* each word has different degrees of saturation on each factor; (c) The factors describing a particular set of words are unique to that set and bear no relation to factors obtained for other sets. In this respect, Deese's conception resembles more closely that of Locke than the modern approaches mentioned. Deese considers the uniqueness of his factors as somehow contradictory to the general factors repeatedly found in the work with the semantic differential and superior to them. It seems obvious, however, that words can and should be described by reference to *both* general and specific factors. The two descriptions complement each other. The description of MUSIC as highly positive on the evaluative factor is perfectly consonant with its description as having a high loading on a specific Factor I.

For a closer evaluation of Deese's analysis, we will turn now to an examination of its face validity, predictive validity, and construct validity. Some factorial structures make good intuitive sense. For instance, in a study of 17 words preselected for their relation to MUSIC (Pp. 76-78), Factor I showed relatively high loadings for *instrument, symphony, sing, note, song, piano, band, horn, opera,* and *music,* and Factor II high loadings for *tone, sound, noise, loud, hear, ear,* and *soft.* It seems that words were split into a "musical" cluster (Factor I) and a "nonmusical" cluster (Factor II). Some of the analyses yield classifications which the reader is willing to accept as sensible and revealing of lexical relations, although he would not have made such classifications *a priori.* The conclusions one can draw from these results about the general validity of this method for studying semantic relations are, however, limited by two considerations. First, these results were obtained for sets of words carefully preselected for their interrelations. An analysis of a random selection of words is more likely than not to produce a chaotic factorial structure. Secondly, even in cases of preselected sets, as are all studies reported in the book, some factorial structures do not separate the words into intuitively acceptable groupings. Face validity can thus be a double-edged sword and Deese is understandably reluctant to depend too heavily on it.

As support for the validity of his measures, Deese cites their ability to predict the results of various verbal learning and behavior tasks. In the absence of systematic comparative analyses, however, it is not possible to assess the advantage of the IC measure and the factor analytic method associated with it over the other word association indices. In fact, in the only comparison reported, a measure of direct interword associations predicts free recall somewhat better than does IC. Thus, the evidence from the verbal learning and behavior laboratory cited in this book, although of considerable interest in itself, can only add support to the general conclusion that free association data can be used to predict the outcome of a variety of verbal tasks.

One may also seek the validity of Deese's analysis in its theoretical underpinning: The author feels that his analysis is grounded in the mediational notion of stimulus-equivalence which asserts that two stimuli, A and B, are equivalent if they share a common response, C. The equivalence in the model case is reflected in the greater ease of learning the association B-D after having learned A-D. In that case, the notion of equivalence is thus specific and fairly well defined. But what kind of equivalence can be attributed to two words connected by a high IC value? And in what sense is the equivalence reflected in one factor different from that reflected in another factor? In various contexts, Deese interprets the idea of equivalence in terms of parts of speech. But such equivalences are certainly not characteristic of the factorial structures where words belonging to different parts of speech turn up indiscriminately in the same factors, unless arbitrarily excluded in the selection of stimulus-words.

Such exclusions are made deliberately in the second part of the book (Chapters 5-8), which deals specifically with the relation between word associations and parts of speech. Using the distinction between paradigmatic responses (same part of speech as the stimuli) and syntagmatic responses (different part of speech), Deese found the following distribution of paradigmatic responses: nouns, 79%; verbs, 52%; adjectives, 50%; and adverbs, 27%. No new or specific proposals are made to account for the syntagmatic responses. Rather, the book concentrates on the analysis of the paradigmatic responses to nouns and adjectives and bases on this analysis two "new laws of association": a law of grouping for nouns and a law of contrast for adjectives. One of the main empirical foundations for the principle of contrast is a study of associations to 278 common adjectives (Pp. 121-128). Of these, Deese isolated 39 pairs of antonyms, such as *alone-together, happy-sad, hard-soft,* and *thick-thin,* which constitute "fundamental contrasts." These pairs have the characteristic of reciprocal association, *i.e.* the two members in each pair tend to elicit one another in a free association test. These pairs are considered fundamental because, it is claimed, each contrast is unique and not reducible to more general factors. The factorial evidence for this claim is scanty and cannot be evaluated without comparable analysis for other words. Moreover, the argument Deese advances that there are sentence-frames in which one contrasting pair could fit well but a seemingly closely related pair could not fit as well, can be made for every word in the language. Every word is in some sense unique (it wouldn't exist if it weren't) and best suited for the expression of a particular idea, but this shouldn't prevent it from sharing common factors with other words.

Deese considers associative reciprocity as a kind of operational definition of antonymy. This is rather surprising in view of the fact that of 100 *S*s, the number of those who responded reciprocally to any given pair never exceeded 40 and was usually much smaller than that. Can one seriously believe that for the *S* who responds to WHITE with *house* rather than with *black,* the two adjectives are not opposites? Or can it be claimed that *inside-outside* which have an average reciprocity score of 40 are two times more opposite than *clean-dirty,* which have a reciprocity score of only 18? One might as well accept the fact that no operational definition is available at present, and may never be forthcoming, for antonymy as well as for other important psychological concepts.

As was already indicated, the 39 antonymous pairs serve to establish the law of contrast for adjectives. The law asserts that "Elements are associatively related when they may be contrasted in some unique and unambiguous way" (P. 165). One may first ask: What is the lexical domain to which this law applies? It appears that the English lexicon was searched through appropriately and only the discussed 39 pairs were found to be both contrastive and reciprocally associated. At best, then, the law of contrast is nothing more than a summary of this finding. It appears that the author himself is uncomfortable with this law for he interprets it in terms of contextual equivalence: contrasts elicit each other because they are used in the same linguistic environments (*e.g.* P. 138). In one study (Pp. 139-140) designed to support this interpretation, a single *S* produced 15 sentences in which he used the words *hot, cold, good,* and *bad.* These adjectives were then deleted and a group of 100 *S*s was asked to fill in appropriate adjectives in the blank spaces. It was found that the opposite was more often substituted in the reconstructed sentences than one of the other two words. For instance, in the sentences where *bad* had originally appeared, *good* was given 4.1% of the cases whereas *hot* was given only 1.9% and *cold* 0.2%. Such studies, however, do not constitute sufficient evidence for the contextual determination hypothesis. To verify this hypothesis, one is under obligation to show that an opposite is more likely to replace its counterpart than *any* other word, not just an arbitrarily selected pair. More generally and importantly, the sentences an *S* gives are not usually reproductions of previously experienced ones, but are new constructions. Therefore, one cannot assume that the sentences an *S* constructs on the request of an experimenter represent a sample of sentences experienced in the past by the *S* or by other speakers in the same community. It is thus unjustified to hold linguistic context responsible for free associative behavior. It appears that free associations are not direct reproductions of past experiences but rather constitute, as does most behavior, productive events related to the past in an indirect and complex manner. At some points, Deese seems to voice a related viewpoint but it is not clear how it relates to the contextual equivalence notion.

The problems of theoretical status and empirical verification raised with respect to the principle of contrast for adjectives apply also to the principle of grouping for nouns. In addition, it suffers from another difficulty. Deese states (P. 165): "elements are associatively related when they may be grouped because they can be described by two or more characteristics in common." It is apparent, however, that every word shares some features with a great many other words, and unless one specifies what features are relevant under what conditions, the principle of grouping remains devoid of explanatory and predictive value. Without such explicit specification the principle of grouping makes little advance over the early philosophical notion of similarity.

Current investigators working with free associations generally exhibit interests that go little beyond the empirical description of functional relations. Against this background, Deese's work stands out in its broad goals and dedicated search for an understanding of the central and most complex problem areas in psychology, language and thinking. The very attempt to deal seriously with these problems in the verbal learning framework, whatever its outcome, must have a profound influence in focusing the attention of investigators in this area on the ultimate goals towards which their research must be directed. It would not be surprising if a re-orientation and clarification of research aims led to a gradual replacement of the free association task in favor of other more explicitly directed and constrained tasks.

MOSHE ANISFELD

Cornell University

The Psychology of Learning. By JOHN F. HALL. Philadelphia and New York, J. B. Lippincott Company, 1966. Pp. vii, 737. $6.95.

This significant text is an accurate reflection of the present state of the psychology of learning. This comment can be taken both as praise and criticism, for the present state of the psychology of learning leaves much to be desired. Fortunately, this book has solid virtues over and above certain apparent ones which may prove to be merely contemporary preferences and its weaknesses, if such they are, reflect upon the existing biases in the field rather than upon the scholarly and competent author.

The functional trend in American psychology has always been strong and it is not surprising that, in this period of reaction to the excesses of the general learning theories of a generation ago, there should be a return to that functionalism. In this spirit, Hall has organized his treatment of learning around a consideration of the variables which affect performance during learning and recall. When theory fails, we reorganize by closely examining the facts; in the area of learning the facts are the empirical relationships obtaining among the variables under study in laboratory situations. Following this orientation, Hall has achieved a consistent and well ordered presentation of the major empirical findings gleaned from classical, instrumental, and verbal learning situations.

Hall views the usual experimental learning situations as "laboratory models from which one may observe the operations of certain variables which contribute to the learning process" (p. 43). Thus in his presentation he makes no distinction among them based upon different learning principles, and often describes the effect of a single variable across a variety of situations. Incidentally, few texts have been so bold as to incorporate the extensive literature on verbal learning into such a general treatment and Hall's thoroughgoing effort to do so may at long last signal the emergence of this previously isolated area to first class membership in the domain of learning.

After describing the common learning situations and methodologies used in laboratories, Hall devotes four chapters to "motivational variables", such as drive, reinforcement, and punishment. There are two chapters on the effects of non-reward, including partial reinforcement and extinction. Variables derived largely but not exclusively from verbal learning studies are treated in two chapters on task variables. These include meaningfulness, instruction, context, etc. Generalization, transfer, retention and forgetting are reviewed in subsequent chapters and, in spite of the otherwise empirical emphasis, one chapter deals with classical theoretical issues—place learning, latent learning, transposition, continuity. The inclusion of the chapter on theory and the pervading recognition that theoretical hypotheses have provided the impetus for much if not most research in learning are evidences that Hall is not anti-theoretical. Like the majority of learning specialists he is keenly aware of the resistance of the great volume of facts to consistent interpretation in terms of any single theory. He chooses, then, to report the facts and to play down theoretical questions concerning them.

The functional approach coupled with the disillusionment with theory will strike a responsive cord with many and will make this a popular advanced undergraduate and graduate text. It is honest, impartial, rigorously explicit. It presents in an orderly manner the evidence that the student of learning should know. This is the way the field appears today and if the book has weaknesses they are those inherent in this way of viewing it.

What might the weaknesses of this particular functional approach include? The major ones seem to derive from the apparent contradictions involved in the rejection of inadequate theories and the continued minute examination of the situations relevant to them. If the theories were inadequate can one accept even the relevance of the "standard" situations to which they give rise? Are laboratory learning situations "models" of anything other than themselves? Will continued detailed study of the traditional situations ever clarify the present chaotic picture? Hall's text, and the field of learning today, give the impression of a revolution half-finished. The old order of theory has been renounced but the everyday business of the laboratory continues unchanged. Nowhere in this text is there the suggestion that new situations, new variables, new conceptualizations may be required to break out of the dilemmas of the past.

The underlying passiveness of the descriptive approach leads to other problems. A readable style is difficult when the subject matter consists primarily of a recitation of the often conflicting empirical results of experiments dealing with a particular variable. It is as though the reader were being presented the evidence concerning some mystery without first being made aware of who was on trial or what the charge was.

Finally, even an "empirical" approach is conceptually hazardous. Hall has recognized the dangers in the assumption that variables themselves can be made comparable across diverse situations. (Are increased numbers of pellets for a rat and increased monetary gain comparable operations to increase reward?) Not so clearly expressed is the fact that without conceptual guides to distinguish learning variables from performance variables no easy way exists to delineate learning as a sub-area of psychology. One consequence in this case is that a large amount of space is devoted to the effects of variables which would ordinarily be considered part of the psychology of motivation.

It is noteworthy that the text contains no effort at a chapter on summary or conclusions. One is abruptly abandoned on page 633, just after the resumé of the data on recall as a function of the number of previous lists learned. This omission emphasizes the unwillingness or inability of learning psychologists at this time to work toward a synthesis of their material which would be relevant for other areas or suggest paths for the future development of learning itself.

These criticisms, it must be recalled, reflect upon the current orientation of the field of learning in general as much as upon this text. However, one may feel the need for change in that orientation, it must be agreed that this is a careful and balanced presentation. (Editorial errors, such as the mislabelling of figures 2.2 and 2.3 [p. 17-18] and the deplorable construction "data . . . is" [p. 47] are rare.) Within its defined objectives, the book does well what it sets out to do.

WALLACE A. RUSSELL
University of Minnesota

Experiments in Induction. By E. B. HUNT, J. MARIN and P. STONE. New York: Academic Press, 1966. Pp. xi, 247. $9.50.

It is a pleasure to quote a dust jacket that describes accurately the contents of a book: "A project is described in which list processing computer programs were used to solve complex logical pattern recognition tasks. The programs and the problems are described in detail. The results of an extensive series of experiments, which included a comparison of simulated and actual human performance on concept learning tasks, are given. The program . . . have also been applied to concept learning tasks in content analysis of text and in medical data processing. The methods of application are outlined."

Part I of *Experiments in Induction* discusses the nature of concept-learning and describes the types of problems used; Part II reports a series of experiments on computer concept-learning systems; Part III and Chapter 9 relate these results to corresponding experiments on human concept-learning; Chapter 8 reports on an application of the computer concept-learning system, in conjunction with the General Inquirer, to content-analysis problems. Miss Marin collaborated in the research of Part III, and Stone in the work of Chapter 8; most of the rest of the research is Hunt's. The concept-learning problems, the computer-simulation programs and the experiments are all described clearly, and in sufficient detail so that they can be replicated.

The eight experiments with computer concept-learning systems (*CLS*) can be viewed either independently, as artificial intelligence approaches to induction, or as a basis for explaining the data from the human concept-learning experiments. All of the *CLS* programs are variants on a basic theme: On the basis of experience with a growing set of items sampled from a universe, "grow" a sorting scheme, in the form of a decision tree, that will classify the items as instances, or not, of the concept known to the experimenter.

A first series of well-designed experiments is concerned with problem-difficulty. Problem-difficulty is greatly affected by the form of the concept, and by the number of relevant dimensions; but only slightly by the number of attribute values and irrelevant dimensions. A second series of experiments explores the effects of modifying the basic *CLS-1* program. (For example, limiting to 30, 20, or even 10 the number of previous items the system can remember decreased its efficacy only slightly, as did allowing it to select new items on the basis of the potential usefulness of the information they contained.)

In general, the lesson that the reader (and the authors) draw from these experiments is that in a "world" having little structure or redundancy, sophisticated inducton schemes will show little advantage over simpler schemes. The result may be stated so as to appear either trivial ("It isn't worth searching for what isn't there"), or paradoxical ("Don't search hard for things hard to find"). It is neither; rather, it is a significant fact about inductive search in the kinds of problem-spaces that have commonly been used in concept learning experiments in the laboratory. It might well suggest to us a Brunswilkian caution in applying the laboratory results to real-world situations (e.g., scientific activity directed towards discovering laws of nature), for the real world may well have far more structure and redundancy than these experimental settings.

A pair of experiments with human subjects, paralleling the simulation experiments, show that concepts of different forms are ordered in roughly the same way by difficulty for humans as for the *CLS*. The *CLS* uniformly makes far fewer errors, and therefore the agreement is not terribly impressive. More impressive is the agreement between simulation and subjects in the specific items on which errors were made. In a final experiment, in the tradition of the binary choice experiments, subjects (and the *CLS*) were asked to find concepts in random data. Here the experiments show some tantalizing but unexplained differences among subgroups of subjects drawn from different academic backgrounds.

The interest of the experimental data rises as the analysis moves gradually from traditional techniques—tests of significance on grouped data—to more detailed reports on the predictability of responses on individual items, and the differences in strategies among different subjects. It is regrettable that the authors stuck as closely as they did to the traditional methods, for they never reached the point (surely requiring protocol-analysis) where they could test in detail in what respects the *processes* incorporated in *CLS* resembled or differed from the *processes* used by human subjects. They are clearly aware of the issues; in this book they do not deal with them.

The chapter dealing with the use of *CLS* in conjunction with the General Inquirer has more a flavor of prospectus than application. The examples cited are rather too simplified and schematic to provide much feel for the potential usefulness of the method.

In the final chapter, and in some earlier passages, the authors relate their research to other work on concept-formation. In general, the discussion is competent, if brief. I must disagree, however, with the assertion that *CLS* is only "superficially similar" to its predecessor *EPAM,* constructed by Feigenbaum. Hunt appears unaware that *EPAM* generally, stores only incomplete stimulus tests and images, hence is capable of generalizing in the manner attributed exclusively to *CLS* on page 192.

Experiments in Induction is a valuable addition to the growing body of literature that combines simulation techniques with human experiments to discover how people discover pattern in the world around them. It provides us with some detailed process models of how computer systems can perform this task. The task ahead is to improve our experimental techniques to the point where we can identify in corresponding detail the processes that account for the human behavior.

Carnegie Institute of Technology HERBERT A. SIMON

Foundations of Physiological Psychology. By RICHARD F. THOMPSON. New York, Harper and Row, 1967. Pp. xxvii, 667. $10.00.

A physiological psychology text must attempt to be all things to all people. It should present a balanced overview of the field, but the author's special interests should lend added zest. The text should be useful for an introductory course in physiological psychology, yet profitable for a graduate student to read, and of interest to the author's peers. *Foundation of Physiological Psychology* comes close to meeting these aims.

This book has a number of particularly desirable aspects. An introductory chapter provides a short but good historical overview of physiological psychology and is followed by two chapters which present the basis of research methods in psychobiology. There is a rather brief presentation of the nature of molecules and chemical bonds, and an excellent discussion of electricity and magnetism. Very good, concise descriptions and explanations of histological procedures for studying both normal and lesioned neural tissue are presented. A valuable discussion of stereotaxic methods is provided, and a clear and fairly detailed description of electrophysiological recording methods.

The introductory and methods section is followed by chapters which present and analyze relevant research. The reader is introduced to a useful and clear overview of neuroanatomy, including a brief account of invertebrate central nervous systems. Action potentials *per se* are reviewed in a satisfactory manner. Synaptic transmission and the subsequent initiation of action potentials are well presented, although an overly dogmatic focus is maintained on the axon hillock as the only possible site of the initiation of action potential. Sensory processes from receptors through thalamic nuclei are presented in a single chapter. Somesthetic responses to skin deformation, heat, joint movement, and hair movement, receive a clear and fairly current treatment. Chemical senses, after an unfortunate initial statement that "Taste proper is classically described as having four qualities . . ." are given an unusually good and complete treatment, which lends no support to the initial four-quality statement. Sensory responses at the cerebral cortex are also treated jointly. Coverage is rather good, although cortical auditory responses to stimuli other than pure tones are neglected.

Two outstanding critical analyses of active research areas are provided. In one, hypothalamus and the limbic system are considered in an integrated fashion. Emotion, food intake, liquid intake, and sexual behavior are clearly and insightfully described in relation to neural structures. The amygdala, the hippocampus, and limbic cortex receive special attention. The second analysis offers an excellent review of the neuronal basis of learning. Biochemical aspects are considered first, with a lucid and detailed discussion of the data on acetylcholine and ribonucleic acid. The possible role of glia is presented, and then the important observations made on invertebrate preparations such as *aphysia* and the thoracic ganglia of the cockroach are described. The controversial experiments on planaria are discussed, and then mammals are specifically considered, both in terms of learning and in regard to memory.

Even in this best of all possible worlds, a perfect text is hard to find. *Foundations of Physiological Psychology* does have its flaws. The attempt to explain the physico-chemical basis of neuronal transmembrane potentials is valiant but unsuccessful. Full responsibility for establishing a potential difference is given to the impermeable organic actions within the cell, while the crucial difference in sodium vs. potassium and chloride permeabilities are ignored. If an electroneutral ion pump (sodium pump) had been introduced here, combined with these permeability differences, a straightforward development of transmembrane potential would have been possible. The situation is only further confused when the Nernst equation is suddenly introduced full grown, with no clear indication of its use in determining equilibrium potentials.

Trouble spots are also present in other neural areas. The presentation of visual photochemistry is many years out of date, while the consideration of the peripheral auditory system presents cochlear structure and function in a disappointing, unoriginal, and dated fashion. Another weak point is the treatment of neural control of movement. The "extrapyramidal system" is not given adequate consideration, while the discussions of the cerebellum and of the vestibular system contain none of the relatively recent observations which have drastically altered our understanding of these structures.

The book has many valuable and well reproduced figures, suggested readings after each chapter, and an excellent bibliography which documents the text in a definitive but unobtrusive manner. The defective portions of the book, while not inconsequential, do not destroy its over-all excellence. This book should be useful in physiological psychology and neurobiology courses at all levels, and many psychologists and physiologists will find it worth reading.

Cornell University

Bruce P. Halpern

Explanation and Human Action. By A. R. LOUCH. Berkeley and Los Angeles, University of California Press, 1966. Pp. viii, 243.

Scientific investigation of human action is impossible: to attempt it is to end in "triviality, redundancy and tautology" (p. 9). This is the main burden of the author's jeremiad. He claims that psychology, considered as "a science with its own explanatory laws, falls in a no-man's land between physiology and the *ad hoc* deliverances of every-day life" (p. 38). Propositions put forward by psychologists are policy recommendations in disguise, serving at most to degrade their hearers into subjects capable only of non-rational behavior. Similar strictures are directed against the social sciences, with copious illustrations of the methodological and conceptual errors of some of its best-known practitioners.

Gainsayers of the possibility of genuinely scientific exploration of man or society usually rely upon the unpredictability of action, and upon men's alleged power to subvert any generalization by acts of free choice. Professor Louch takes the more radical line of claiming that the very concepts needed to identify instances of action are "moral", not neutrally descriptive. He holds that the use of such concepts as *action, desire, motive, goal, intention* and *purpose,* demands observation, to be sure, but also a crucial assessment of the value of possible outcomes: relevant features of the actions discussed must be appraised from the agent's stand-point as good or bad, appropriate or inappropriate, desirable or undesirable. Consequently, Louch will accept as relevant explanations of human action only those that by exhibiting the agent's effective reasons reveal how he was "justified" or "entitled" to do what he did. (Louch tends to ignore cases where the explainer must appear in the role of critic rather than counsel for the defense.)

But why should reason-providing-explanations embarrass even the most positivistic of psychologists? A woman refuses to walk across a ploughed field, because she is afraid of the snakes lurking in the furrows: this seems to be a clear case of the "*ad hoc* deliverances of common sense" that Louch endorses. No doubt, this tale, if true, does characterise the action as reasonable in a weak sense, and so in a way "justifies" what was done. Yet the agent's having the beliefs that she did and arguing, correctly or incorrectly, as she did, are still matters of fact, calling for value-free observation, without recourse to special appraisal or assessment. Endorsing her reasons would, of course, be another matter, but Louch is careful to emphasize that he is committed to no such thing. In such cases as this, Louch seems guilty of a fairly obvious confusion between the agent's own use of a "moral concept", which may well involve assessment, and the onlooker's use of the same concept as part of a neutral description of fact.

Louch thinks such reason-providing-explanations are peculiarly "context-bound" and resistant to generalization: a commonsense explanation of one action, appropriate to a particular setting, may fail to fit another acion that is at least superficially similar. The best I can make of this is the view, acceptable enough, that commonsense is usually satisfied with sufficient, rather than with necessary, reasons for problematic actions. But even so, generalization is not precluded. John Doe moved his queen because it was in danger and he wanted to save it (and so on); but then so would any competent chessplayer in a like case. (No tautology lurks here, since we have independent tests for competence at the game.)

Louch's intemperate dismissal of the social sciences would acquire some color of plausibility if experimenters and theorists still treated Pearson's *Grammar of*

Science as gospel. But Louch's imputation of such naive positivism seems to have little foundation.

The author's style is lively, even entertaining, but the general tendency of his exposition seems to me to be obscurantist. He is too impressed by the virtues of common-sense explanation and too little conscious of its obvious limitations. As soon as reasons for action have been provided, an enquiring mind will want to press on to questions about the provenance and etiology of such reasons, thereby leaving the realm where rationality, even of the qualified sort that Louch has in mind, still reigns. I cannot see that Louch has provided good reasons for thinking such enquiries unprofitable or impossible. But then it is uncommonly hard to discover his own reasons for his extreme conclusions, since he is too often satisfied with rhetoric, where some careful argument would have been welcome.

Cornell University MAX BLACK

Molecular Psychobiology: A Chemical Approach to Learning and Other Behavior. By JOHN GAITO. Springfield, Illinois, Charles C Thomas, 1966. Pp. xix, 259. $9.50.

In recent years, an increasing number of psychologists have ventured into the field of biochemistry to learn facts, principles, and techniques that they might use in investigating bodily mechanisms of behavior. This text might be seen as a traveler's diary and guide by one such psychologist—with a basic atlas and description of terrain already covered, accounts of explorations in progress, prospectuses for future trips, and even a section on regional cookery (the appendices on procedures). The book has both the assets and liabilities of such a guide, since it strives to give an up-to-date treatment of a rapidly changing area that is much in the news, rather than attempting a finished work.

The author, Dr. John Gaito, has been occupied with possible biochemical mechanisms of learning and memory since 1960. He has conducted research on this topic at Kansas State University and, since 1964, at York University in Toronto.

The book seems to have been aimed at several different audiences, with a resulting unevenness and lack of focus, especially in the earlier chapters. The foreword by the editor of the Living Chemistry Series states that the series was conceived "to advance the newer knowledge of chemical medicine in the cause of clinical practice," but there seems to be little content of direct clinical relevance. Chapter 2 is a 25-page account of "Behavior and the Nervous System" that is too general or elementary to be of help to either a physician or a psychologist, although it may be useful to a physical scientist getting started in this field.

The audience that can probably gain most from Gaito's book includes other researchers in memory mechanisms and also students of physiological psychology. Such students will find Section II a helpful basic treatment of the role of nucleic acids in genetic mechanisms. This section is written at a more intensive and complex level than are the preceding sections on basic neural functions and learning.

The heart of the book is Section III which seeks to relate behavioral events—chiefly learning and memory—to neurochemical events—chiefly those directly involving *DNA* and *RNA*. The key question here is whether memory may be stored in terms of altered structure of either *DNA* or *RNA*. Holger Hydén started this sort of speculation in 1958 with his hypothesis that patterns of arriving impulses could modify the RNA in neuron and thus "tune" it permanently to that pattern. Gaito surveys a number of studies by Hydén and others meant to test this hypothesis.

He concludes "that there is no crucial experiment which has unequivocally implicated *RNA* uniquely in memory functions" (p. 165). Other *DNA-RNA* memory hypotheses are also discussed and evaluated.

To stimulate further advances, Gaito sets forth designs for some research projects that he considers important and feasible as next steps to be taken. These will clearly be of interest to those active in the field, and the reasoning involved will help students understand the issues.

The final chapter is an ambitious attempt to state what can be said about the physiological and biochemical mechanisms of memory and to suggest speculations where facts are lacking. This chapter provides a revealing indication of current knowledge and gaps in this field.

The appendices on biological and chemical methods may be helpful to the physician or psychologist who wants to have an idea of the techniques and problems involved. Since methods in this field are evolving rapidly, those described here will not necessarily long remain current, and they should therefore be taken only as examples.

One of the assets of the book is a list of almost 400 references which should be of considerable help to those who wish further access to this field. Very few of the items are dated later than 1964. There is also a useful index. The book deserves better editing than it received, for typographical errors and ungrammatical expressions frequently distract the reader.

University of California, Berkeley MARK R. ROSENZWEIG

Fortran Programming for the Behavioral Sciences. By DONALD J. VELDMAN. New York, Holt, Rinehart and Winston, 1967. Pp. x, 406. $5.95.

This is one of the new breed of books in which *FORTRAN* programs and worked examples are presented in detail. To the researcher who likes to do his own *FORTRAN* programming or at least likes to be able to read what goes on during the execution of a *FORTRAN* run such books make fine friends. However, the novice is likely to be overwhelmed by chapters on how a computer works and how a programming language may be used until such time as he *actually does know* how a computer works and how a programming language is used. This is an axiom true of all technical manuals. Inevitably, success depends upon getting ones hands on a computer and developing useful contacts at the computation center.

Veldman covers a lot of ground in this book—and it is of uneven difficulty. Chapters 1-4 and 6 cover the fundamentals of computer systems and the *FORTRAN* language. These chapters will have to be read more than once by the novice. As he becomes more familiar with the way things are at a computation center he will appreciate Chapters 1-3 more. As he learns to read and write *FORTRAN* programs he will appreciate Chapter 4 and especially Chapter 6, which is a much clearer account of the use of functions and subroutines than he is likely to find in technical manuals. For the beginner Chapter 5 is an excellent one. Twelve *FORTRAN* programs, graded in difficulty, are presented with truly commendable clarity. Chapter 7, covering fundamental matrix operations, is brief but clear.

The remaining chapters are certainly not intended for the neophyte since they contain methodologies of considerable sophistication and do not include even the barest mathematical development in most cases. Chapter 8 deals with programs for scoring and interpreting multiple-choice tests, questionnaires and rating scales.

Chapter 9 contain a good program for performing factor analysis and varimax rotations. Chapter 10 deals with some simple *ANOVA* designs and multiple discriminant analysis. Chapter 11 deals with canonical correlation and regression. Chapter 12 includes an interesting clustering program, a successive interval scaling program (which could be improved), and a chi-square test for contingency tables. Finally, Chapter 13 deals with the analysis of verbal data and this is an important chapter. It handles a problem of considerable importance to psychologists, it points up the fact that *FORTRAN* can be used successively for the analysis of open-ended questions, and it is very clearly written.

In comparison with other books of this type (notably, W. W. Cooley and P. R. Lohnes, *Multivariate Procedures for the Behavioral Sciences,* 1962—to which Veldman makes repeated reference), Veldman's book stands up well. It offers a useful package of FORTRAN programs of the more standard variety and also a number of not-so-standard programs, including *PRBF* which computes the exact probabilities of *t*-values, *F*-ratios and X^2-values; *HGROUP* which does a hierarchical cluster analysis in a straightforward manner; and *LISTB* which performs a compilation and analysis of verbal data. If each psychology department could develop a package of utility programs such as suggested in this book, then analysis of data would not be as time-consuming as it often is. As Veldman cautions, however, "the machine's role should be that of an extremely accurate and efficient research assistant" (P. 191) and the availability of useful programs should not give the user licence to practice without going through the discipline of understanding the methodologies themselves. This is a book for research workers who know their methodology and who wish to adopt *FORTRAN* and the computer as their assistants—and it is a good, well-edited book for that purpose.

Dartmouth College

VICTOR E. MCGEE

The Nature of Psychological Inquiry. By RAY HYMAN, Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. xi, 116. Cloth $3.95; Paper, $1.50.

This book attempts to describe in an elementary way how the psychologist as scientist proceeds in the conduct of enquiry. The approach is behavioral rather than philosophical; that is, the author's concern is more with the behavior of the psychologist as scientist, rather than with any particular philosophy of science. The approach is naturalistic, *e.g.* What is the psychologist actually perceived to do in the conduct of enquiry? This level of discourse leaves ample room for both agreement and disagreement between the author and the reader.

The author traces the progress of enquiry through five stages, these being concerned with getting ideas, gathering facts, processing data, interpreting data, and communicating results. Although these titles have the dull ring of the pedestrian textbook, Hyman's book is neither dull nor pedestrian. Quite to the contrary, it is provocative, thoughtful, well written, and, indeed, at times original. It is worth reading for itself, apart from its textbook role. The discussion of the different stages of enquiry are well documented with examples of particular investigations. These examples are skillfully integrated into the text to illustrate clearly the points the author wishes to make.

The book is in the Prentice-Hall Foundations of Modern Psychology Series. It is intended for use at the introductory level along with others in the series. Apart from its use in teaching psychology at the introductory level, it will probably enjoy success as a supplementary text in courses in experimental psychology. It will, I think, prove particularly useful as a basis for discussion in seminar groups.

While respecting the author's point of view, I found points of disagreement more numerous in the chapter "Processing the Data," than elsewhere. The author is critical of statistical method. He writes "Within recent years psychologists have become disenchanted with the ability of statistical methods automatically to further their research." He points to "The ever increasing number of studies in which uncritical use of statistical procedures seems to obscure discovery . . . to the tendency of many investigators to select their problems to fit existing statistical methods rather than the other way around." These are conventional criticisms of statistical method, and everyone would agree to the presence of instances where they are justified. Despite this, I do not think that these are substantive criticisms in this year 1966. Psychologists are at present very sophisticated in the appropriate uses of statistics. A further point is that research in psychology, and elsewhere, is not independent of methods, statistical or otherwise, instruments, theories, and access to human and animal subjects. All these exert a directional force on research. They dictate in large measure what problem the investigator selects, and what he is able to do with it. The point here is that statistical method cannot be singled out for special consideration. If a problem exists it exists at a more general level—methods, instruments, theories, subjects, animals and computers. The negative influence, if any, of statistics on research is a particular instance only.

Although I could challenge the author on a number of other points, I think he has written a good book. I enjoyed reading it.

McGill University GEORGE A. FERGUSON

Amines and Schizophrenia. Edited by HAROLD E. HIMWICH, SEYMOUR S. KETTY, and JOHN R. SMYTHIES. New York, Pergamon. 1967. Pp. ix, 290. $11.00.

Occasionally a symposium comes along which marks a decided step toward the solution of a problem though not the solution itself. The one under review is exactly that. For many years the hope that a biochemical basis for psychosis would be found was supported by a number of biochemical clues that seemed to be pointing in a common direction. All the tranquilizers, antidepressants and hallucinogenic drugs (many discovered by accident and through no common thread) contained a few common chemical features. An indole nucleus was one, a so-called LSD-tail was another, and the presence of methyl groups in hallucinogenic drugs was a third. What they meant was vague and tantalizing but essentially unanswerable. Though no answer can yet be given, the common direction is now so strong that one can almost taste a rush to the goal line. In recent years, Friedgood and co-workers at N.Y.U. reported that 3,4-dimethoxyphenylethylamine, a derivative of dopamine, the amine in the extrapyramidal system, is found in schizophrenics but not normals. His clue: tranquilizers usually produce extrapyramidal side effects. Perry has failed to find it in schizophrenics on a plant-free diet. Is schizophrenia, as Perry suggests, an inability to handle normal dietary components (P. 66)? The methyl-problem implicates methyl-donors and it was reported in 1961 that methionine loading schizophrenic patients produces a toxic psychosis and the tantalizing thought is that it may be an actual exacerbation of the existing psychosis (Kety, p. 126). Tryptophan loading produced a relaxation of these patients. Richter suggests that a disorder of tryptophan metabolism may be involved in psychotic depression. The symposium starts with an excellent brief overview by Smythies who delineates three hypotheses of schizophrenia, biochemically considered. All through the paper tantalizing hints of all these recur—the catechol amine hypothesis, the tryptamine hypothesis, the transmethylation hypothesis. As if to nail down the clues in recent days the new hallucinogen, called *STP* in the papers, turns out to be a triply-methylated phenylethylamine. The things that do the kicking are evidently very small in number. These are only a few of the goodies. Himwich has followed the action of methylated amines neurophysiologically in the rabbit. Friedgood reports that schizophrenic liver methylates dopamine. Jenner reports changes in periodic catatonia. Kety's summarizing comments (Pp. 271-277) are frank in stating the weaknesses of all hypotheses under discussion while stressing the advances made with them. Time marches on. It still remains terribly difficult to do controlled research on mental illness. Control problems repeatedly plague the interpretation of these intriguing findings. But at last the hypotheses have been winnowed down. Many good teams are working them over. We are going to have to know more about psychiatric classification before the problem is cracked, but as this symposium clearly attests, the confusion appears to be over. The hypotheses may all turn out to be wrong, but there is a sense of excitement in the air.

Yeshiva University ALLAN C. GOLDSTEIN

Choice, Strategy, and Utility. By SIDNEY SIEGEL in collaboration with ALBERTA ENGVALL SIEGAL and JULIA MCMICHAEL ANDREWS. New York, McGraw-Hill, 1964. Pp. xii, 180. $5.95.

If you are interested in the behavior of people who must repeatedly predict which of two or three events will occur on the next trial, then you should read this little book. It reports six experiments which test the authors' model of repetitive choice-behavior, a model that is surprisingly successful in predicting group averages of individual data. The model does well primarily because it emphasizes the decision-theoretic aspects of choice-behavior, and so considers not only the event-probabilities but also two types of utilities: the utility of a correct choice and the utility of varying one's choices (the negative utility of boredom, so say the authors). The ratio of these two utilities is estimated from the choice-probabilities, averaged over subjects, obtained under one set of event-probabilities. The utility-ratio is then used to predict the average choice-probabilities for data obtained under a different set of event-probabilities. For the experiments that made quantitative tests of the model, the median discrepancy between predicted and obtained choice-probabilities was only 0.009 and the difference rarely exceeded 0.020.

The book has two major shortcomings. In the first place, although the experimental paradigm used is the same as that found in studies of probability-learning, no mention of probability-learning is ever made. This is a pity, for the book has at least one very important implication for probability-learning: a subject who does not always maximize the number of correct choices may behave this way simply to alleviate boredom. The many experiments in the literature that attribute subjects' non-optimal performance to various task-variables, including properties of the stimulus-sequence, are quite likely showing that bored, unmotivated subjects will seize on any aspect of the task that they can use to reduce their boredom. Seen in this light, simple repetitive choice-tasks might better be classified under the heading "psychology of boredom" than "probability-learning." Secondly, the authors give little in the way of real discussion, integration, suggestions, or implications. I was left with the feeling that the experiments were too neat, the results too tidy, the conclusions too simple. For example, the authors are concerned throughout the book with only group data. Would the utility-parameters estimated for one person predict that same person's behavior in another experiment?

Undoubtedly, the depth missing in this book would have been provided by the author had he lived to complete it. Instead, the co-authors, his wife and a colleague, have felt it necessary to restrict the scope of the book. What they have presented, they have presented clearly and well.

Brunel University,
London, England

LAWRENCE D. PHILLIPS

Sensory Restriction: Effects on Behavior. By DUANE P. SCHULTZ. New York, Academic Press, 1965. Pp. viii, 216. $7.50.

This monograph is a comprehensive handbook on effects of sustained restriction of sensory or perceptual input. In a dozen years, the quantity of pertinent research studies has been remarkable. This review is "an attempt to provide empirical support for the proposition that man needs constantly varying forms of stimulation to function adaptively in his environment" (p. 1), a theme which has already generated much advocacy and criticism.

The text is clear and readable although the style is often too unvarying to be stimulating. It is a scholarly compendium, emphasizing coverage and brevity. The author is frequently content with confining an apparently fruitful interpretation to a single sentence. The paucity of his conclusions presumably reflects a quite understandable view that the diversity in the literature supports only a few patterns or trends. While the contributions of some factors, such as immobility and set or suggestion, are discussed carefully, some difficult but crucial problems seem slighted. Precisely what are the ranges and types of stimulation available to subjects in each of the experimental environments created by the various laboratories? Since both perceptual deprivation and sensory deprivation involve very low levels of patterning of stimulation, why is the greater amount of sensory input in perceptual deprivation associated with more impairment than that observed in the other condition?

The author's colleagues will wish he had been able to be more integrative and more critical. Rarely does he note deficiencies in design or interpretation as he condenses a published report. There is little heuristic novelty in his reconceptualization of the need for stimulation as "sensoristasis", "a drive state of cortical arousal which impels the organism (in a waking state) to strive to maintain an optimal level of sensory variation" (p. 30). Seven "predictions" from his model are offered without elaboration, more on an intuitive than a deductive basis. Even though the book could profitably have been longer, it is still an excellent and up-to-date summary of the topic which will be valuable to many people, especially to inquiring graduate students.

University of Chicago

DONALD W. FISKE

Bases for Effective Reading. By MILES A. TINKER. Minneapolis: University of Minnesota Press, 1965. Pp. vi, 322, $7.50.

In this book, Miles A. Tinker, Professor Emeritus, University of Minnesota, brings together for research workers and practitioners in the field, the findings and conclusions from hundreds of systematic studies dealing with fundamental aspects of the reading process. Much of the research reported is his own, carried out during his thirty-two years at the University of Minnesota.

The book is divided into five parts: Part I contains separate chapters on the research on perception in reading among adults and children, with an attempt to highlight the similarities and differences. It concludes with a separate chapter on comprehension in reading. Part II presents a systematic review of the research on eye movements, together with an interpretation of the findings for the improvement of reading instruction. Part III reviews the research on scientific typography; Part IV, visual functions and illumination; and Part V, appraisal of reading proficiency. In all of the latter sections, the research findings are interpreted in terms that are useful for practical decision making by book publishers, teachers, and others concerned with reading efficiency and improvement of reading.

This is an extremely useful book not only for practitioners but for research workers. Its strength lies in bringing together the research findings on many diverse aspects of the reading process—aspects that have recently disappeared from comprehensive works in the field of reading. The research worker also becomes aware of the large areas in reading that are still unexplored or inadequately explored—where the author cites current practices and expert opinion in place of research findings to substantiate some of his generalizations.

Harvard University

JEANNE CHALL

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review).

ALPERN, MATHEW; LAWRENCE, MERLE; and WOLSK, DAVID, *Sensory Processes.* Belmont, Calif., Brooks/Cole Pub. Co., 1967. Pp. vii, 151. Trade, $2.35; Text, $1.75 (paperback).

ARGYLE, MICHAEL, *The Psychology of Interpersonal Behaviour.* Baltimore, Md., Penguin Books, 1967. Pp. 223. $1.25 (paperback).

BARNETT, S. A., *The Rat. A Study in Behaviour.* Chicago, Ill., Aldine Pub. Co., 1966 (first published 1963). Pp. xvi, 228. $2.95 (paperback).

BAUMEISTER, ALFRED A. (ed.), *Mental Retardation. Appraisal, Education, and Rehabilitation.* Chicago, Ill., Aldine Pub. Co., 1967. Pp. viii, 419. $12.50.

BERRILL, N. J., *Man's Energetic Mind. Man's Progress through Time—Trees, Ice, Flood, Atoms and the Universe.* New York, Dodd, Mead & Co., 1955. Pp. xii, 308. $1.95 (paperback).

BORING, EDWIN G., and LINDZEY, GARDNER (eds.), *A History of Psychology in Autobiography.* Vol. V. New York, Appleton-Century-Crofts, 1967. Pp. xiii, 449. $8.00.

BOTWINICK, JACK, *Cognitive Processes in Maturity and Old Age.* New York, Springer Pub. Co., 1967. Pp. x, 212. $5.75.

BRUSSEL, JAMES A., and CANTZLAAR, GEORGE L., *The Layman's Dictionary of Psychiatry.* New York, Barnes & Noble, 1967. Pp. xiv, 269. Cloth $4.50; Paper, $1.95.

BUGENTHAL, JAMES F. T. and others, *Challenges of Humanistic Psychology.* New York, McGraw-Hill, 1967. Cloth, $6.95; Paper, $4.95.

CHASE, CLINTON I., *Elementary Statistical Procedures*. New York, McGraw-Hill, 1967. Pp. vii, 245. $6.95.

CHASSAN, J. B., *Research Design in Clinical Psychology and Psychiatry*. New York, Appleton-Century-Crofts, 1967. Pp. xviii, 280. $5.75.

COLLOQUES INTERNATIONAUX DU CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, *Les Modèles et la Formalisation du Comportement*. Paris, Centre National de la Recherche Scientifique, 1967. Pp. 424. 50 F.

COPEL, SIDNEY L., *Psychodiagnostic Study of Children and Adolescents*. Springfield, Ill., C. C. Thomas, 1967. Pp. xii, 201. $7.75.

DALE, D. M. C., *Deaf Children at Home and at School*. London, Eng., Univ. of London Press, Ltd., 1967. Pp. 272. 35/-

DANCE, FRANK E. X. (ed.), *Human Communication Theory. Original Essays*. New York, Holt, Rinehart and Winston, 1967. Pp. x, 332. $6.95.

DARLEY, FREDERIC L. (ed.), *Brain Mechanisms Underlying Speech and Language. Proceedings of a Conference Held at Princeton University, Nov. 9-12, 1965*. New York and London, Grune & Stratton, 1967. Pp. x, 261.

DE MONTMOLLIN, MAURICE, *Les Systèmes Hommes-Machines*. Paris, Presses Universitaires de France, 1967. Pp. 248 (paperback).

DUMAZEDIER, JOFFRE (Foreword by DAVID RIESMAN, trans. by STEWART E. MCCLURE), *Toward a Society of Leisure*, New York, Free Press, 1967. Pp. xi, 307. $6.95.

EATON, MERRILL T., and PETERSON, MARGARET, H., *Psychiatry. Medical Outline Series*. Flushing, N. Y., Medical Examination Pub. Co., 1967. Pp. 564. $6.00 (paperback)

EISENBERG, RAFAEL, *The East West Conflict. Psychological Origin and Resolution*. New York, Diplomatic Press, Inc., 1967. Pp. 231. $4.95.

ELINSON, JACK; PADILLA, ELENA; and PERKINS, MARVIN E., *Public Image of Mental Health Services. Report of a Joint Project of the Columbia University School of Public Health and Administrative Medicine and the New York City Community Mental Health Board*. New York, Mental Health Materials Center, 1967. Pp. xi, 288. $7.00.

ELZEY, FREEMAN F., *A First Reader in Statistics*. Belmont, Calif., Wadsworth Pub. Co., 1967. Pp. 71. $1.25 (paperback).

FENTON, NORMAN; REIMER, ERNEST G.; and WILMER, HARRY A. (eds.), *The Correctional Community. An Introduction and Guide*. Berkeley and Los Angeles, Univ. of California Press, 1967. Pp. xi, 119. $2.25 (paperback).

FITTS, PAUL M., and POSNER, MICHAEL I., *Human Performance*. Belmont, Calif., Brooks/Cole Pub. Co., 1967. Pp. x, 162. Trade, $2.35; Text, $1.55. (paperback).

FREEMAN, HARROP A., *Counseling in the United States*. Dobbs Ferry, N. Y., Oceana Publ., Inc., 1967. Pp. vi, 322. $10.00.

FRIEDMAN, NEIL, *The Social Nature of Psychological Research: The Psychological Experiment as a Social Interaction*. New York, Basic Books, 1967. Pp. xiv, 204. $6.95.

GLASER, BARNEY G., and STRAUSS, ANSELM L., *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, Ill., Aldine Pub. Co., 1967. Pp. x, 271. $6.75.

GOLDMAN, GEORGE S., and SHAPIRO, DANIEL (eds.), *Developments in Psychoanalysis at Columbia University. Proceedings of the 20th Anniversary Conference. Columbia Psychoanalytic Clinic for Training and Research, October 1965*. New York, Hafner Pub. Co., 1966. $12.50.

GOWAN, JOHN CURTIS; DEMOS, GEORGE D.; and TORRANCE, E. PAUL, *Creativity: Its Educational Implications*. New York, Wiley, 1967. Pp. xii, 336. Cloth, $7.95; Paper, $4.95.

HALL, JOHN F. (ed.), *Readings in the Psychology of Learning*. Philadelphia, Pa., Lippincott, 1967. Pp. xi, 525. $4.25 (paperback).

HARDING, ROSAMOND, *An Anatomy of Inspiration*. With an *Appendix* on *The Birth of a Poem* by Robert Nichols. (2nd ed.) New York, Barnes & Noble, 1967. Pp. xvii, 145. $6.25.

Hays, William I., *Quantification in Psychology.* Belmont, Calif., Brooks/Cole Pub. Co., 1967. Pp. viii, 87. Trade, $2.35; Text, $1.75 (paperback).

Hellmuth, Jerome (ed.), *Disadvantaged Child. Vol. 1.* Seattle, Wash., Seattle Seguin School, Inc., 1967. Pp. 499. $10.00.

Hilgard, Ernest R., and Atkinson, Richard C., *Introduction to Psychology.* 4th ed. New York, Harcourt, Brace and World, 1967. Pp. xv, 686. $8.95.

Hillebrand, M. J., *Psychologie des Lernens und Lehrens. Eine anthropologisch-psychologische Grundlegung.* Bern und Stuttgart, Verlag Hans Huber, 1967. Pp. 217. Fr./DM 23.—

Hinton, John, *Dying.* Baltimore, Md., Penguin Books, 1967. Pp. 208. 95c (paperback).

Hirschi, Travis, and Selvin, Hanan C., *Delinquency Research. An Appraisal of Analytic Methods.* New York, Free Press, 1967. Pp. xiv, 280. $7.95.

Isaacson, Robert L.; Hutt, Max L.; and Blum, Milton L. (eds.), *The Science of Behavior; The Science of Interpersonal Behavior.* New York, Harper & Row, 1967 (first published 1965, 1966). Pp. xv, 815. $10.00.

Johnson, Doris J., and Myklebust, Helmer R., *Learning Disabilities. Educational Principles and Practices.* New York and London, Grune & Stratton, 1967. Pp. xii, 336. $9.75.

Kimble, Gregory A. (ed.), *Foundations of Conditioning and Learning.* New York, Appleton-Century-Crofts, 1967. Pp. xii, 696. $8.75.

Kleinmuntz, Benjamin (ed.), *Concepts and the Structure of Memory.* New York, Wiley, 1967. Pp. xiv, 286. $7.95.

Kochen, Manfred (ed.), *The Growth of Knowledge. Readings on Organization and Retrieval of Information.* New York, Wiley, 1967. Pp. xvii, 394. $14.95.

Kretchmer, Ernst, *Der sensitive Beziehungswahn. Ein Beitrag zur Paranoiafrage und zur psychiatrischen Charakterlehre.* (4th ed.) Berlin and New York, Springer-Verlag and Heidelberg, 1966. Pp. v, 236. DM 36—.

Laurie, Peter, *Drugs. Medical, Psychological, and Social Facts.* Baltimore, Md., Penguin Books, 1967. Pp. 174. 95c (paperback).

Lehiste, Ilse (ed.), *Readings in Acoustic Phonetics.* Cambridge, Mass., M.I.T. Press, 1967. Pp. ix, 358. $10.00.

Lowenfeld, Margaret, *Play in Childhood.* New York, Wiley, 1967. Pp. 341. $2.25 (paperback).

Lowry, Thomas P. (ed.), *Hyperventilation and Hysteria. The Physiology and Psychology of Overbreathing and Its Relationship to the Mind-body Problem.* Springfield, Ill., C. C. Thomas, 1967. Pp. xi, 192. $8.75.

Lynd, Helen Merrell, *Toward Discovery. Curved and Rendered by Reason.* New York, Hobbs, Dorman & Co., 1967 (first published 1965). Pp. xiii, 275. $6.50.

Mahrer, Alvin R. (ed.), *The Goals of Psychotherapy.* New York, Appleton-Century-Crofts, 1967. Pp. viii, 301. $6.00.

Mann, Richard D. (in collaboration with Graham S. Gibbard and John J. Hartman), *Interpersonal Styles and Group Development. An Analysis of the Member-Leader Relationship.* New York, Wiley, 1967. Pp. ix, 308. $8.25.

Masterson, James F., Jr., *The Psychiatric Dilemma of Adolescence.* Boston, Mass., Little, Brown & Co., 1967. Pp. xiv, 217. $9.00.

Meredith, William, *Basic Mathematical and Statistical Tables for Psychology and Education.* New York, McGraw-Hill, 1967. Pp. ix, 333. $3.95.

Millenson, J. R., *Principles of Behavioral Analysis.* New York, Macmillan, 1967. Pp. xxii, 488. $7.95.

Mills, C. Wright, *The Sociological Imagination.* New York, Oxford Univ. Press, 1967 (first published 1959). Pp. 234. $1.75 (paperback).

Newman, Ruth G., *Psychological Consultation in the Schools. A Catalyst for Learning.* New York, Basic Books, 1967. Pp. xv, 300. $6.95.

Rapaport, David (ed. by Merton M. Gill), *The Collected Papers of David Rapaport.* New York, Basic Books, 1967. Pp. x, 943. $17.50.

ROSOW, IRVING, *Social Integration of the Aged.* New York, Free Press, 1967. Pp. xix, 354. $8.95.

RULON, PHILLIP J., TIEDMAN, DAVID V., TATSUOKA, MAURICE M., and LANGMUIR, CHARLES R., *Multivariate Statistics for Personnel Classification.* New York, Wiley, 1967. Pp. xi, 406. $12.95.

SACERDOTE, PAUL, *Induced Dreams. About the Theory and Therapeutic Applications of Dreams Hypnotically Induced.* New York, Vantage Press, 1967. Pp. x, 174. $4.00.

SANDERS, A. F. (ed.), *Attention and Performance.* Amsterdam, North-Holland Pub. Co., 1967. Pp. xi, 452. $16.50 (paperback).

SARASON, SEYMOUR B., DAVIDSON, KENNETH S., LIGHTHALL, FREDERICK F., WAITE, RICHARD R., and RUEBUSH, BRITTON K., *Anxiety in Elementary School Children. A Report of Research.* New York, Wiley, 1967 (first published 1960). Pp. viii, 351. $2.65 (paperback).

SARTAIN, AARON QUINN, NORTH, ALVIN JOHN, STRANGE, JACK ROY, and CHAPMAN, HAROLD MARTIN, *Psychology: Understanding Human Behavior.* 3rd ed. New York, McGraw-Hill, 1967. Pp. xi, 528. $8.50.

SCHON, DONALD A., *Invention and the Evolution of Ideas.* New York, Barnes & Noble, 1967 (first published 1963). Pp. xiii, 208. Cloth, $5.75; Paper, $4.00.

SHERIF, CAROLYN W., and SHERIF, MUZAFER (eds.), *Attitude, Ego-Involvement, and Change.* New York, Wiley, 1967. Pp. x, 316. $9.75.

SLUCKIN, W., *Imprinting and Early Learning.* Chicago, Ill., Aldine Pub. Co., 1967 (first published 1965). Pp. x, 147. $2.45 (paperback).

SPENCE, KENNETH W., and SPENCE, JANET TAYLOR (eds.), *The Psychology of Learning and Motivation. Advances in Research and Theory. Vol. 1.* New York, Academic Press, 1967. Pp. x, 381. $11.50.

TALBOT, TOBY (ed.), *The World of the Child. Birth to Adolescence from the Child's Viewpoint.* Garden City, N.Y., Doubleday & Co., 1967. Pp. xi, 468. $6.95.

TEEVAN, RICHARD C., and SMITH, BARRY D., *Motivation.* New York, McGraw-Hill, 1967. Pp. xix, 172. Cloth, $4.95; Paper, $2.50.

TURNER, MERLE B., *Philosophy and the Science of Behavior.* New York, Appleton-Century-Crofts, 1967. Pp. xvii, 539. $9.50.

VON BÉKÉSY, GEORG, *Sensory Inhibition.* Princeton, N. J., Princeton Univ. Press, 1967. Pp. x, 265. $8.50.

WALKER, EDWARD L., *Conditioning and Instrumental Learning.* Belmont, Calif., Brooks/Cole Pub. Co., 1967. Pp. xii, 161 (paperback).

WALKER, EDWARD L., and MCKEACHIE, WILBERT J., *Some Thoughts about Teaching the Beginning Course in Psychology.* Belmont, Calif., Brooks/Cole Pub. Co., 1967. Pp. viii, 84. Trade, $2.35; Text, $1.75 (paperback).

WILENSKY, HAROLD L., *Organizational Intelligence. Knowledge and Policy in Government and Industry.* New York, Basic Books, 1967. Pp. xiv, 226. $5.95.

WOLFGANG, MARVIN E., and FERRACUTI, FRANCO, *The Subculture of Violence. Towards an Integrated Theory of Criminology.* New York, Barnes & Noble, 1967. Pp. xxiii, 387. Cloth, $11.50; Paper, $5.75.

ZAJONC, R. B. (Trans. by Y. NOIZET, Preface by C. FLAMENT), *Psychologie Sociale Experimentale.* Paris, Dunod, 1967. Pp. x, 139 (paperback).

ZUBIN, JOSEPH and HUNT, HOWARD F. (eds.), *Comparative Psychopathology: Animal and Human. Proceedings of 55th Annual Meeting of American Psychopathological Association, 1965.* New York and London, Grune & Stratton, 1967. Pp. xv, 350. $15.75.

ZUCKER, HERBERT, *Problems of Psychotherapy.* New York, Free Press, 1967. Pp. viii, 194. $5.95.

INDEX

By RAND B. EVANS, Wright State University

(The names of authors of original articles are printed in CAPITALS and Small Capitals; of authors of books reviewed in roman, and of reviewers in *italics*.)

ALPERSON, B. 416
Alpert, R. 481
ANDERSON, J. P. .. 643
Anisfeld, M. 646
ANSTIS, S. M. ...51, 572
Appley, M. H. 319
ATKINSON, J. 572
Atkinson, J. W. ... 319
Atkinson, R. C. ... 482

BAKAN, P. 416
BECK, J. 59
Beck, J. 159
Beritoff, J. S. 478
Bernbach, H. 482
BEVAN, W. 309
BIRNBAUM, I. M. .. 363
BITTERMAN, M. E. . 269
Black, M. 655
BLANTON, R. L. ... 391
BLESSING, W. W. .. 250
BOITANO, J. J. 73
Boring, E. G. 150
Bower, G. H. 482
BRENT, S. B. 133
BRITTEN, R. 640
Brown, R. 324
BROWNING, D. 466
BURNHAM, C. A. .. 434
Bush, R. R. 322

CAMPBELL, D. 438
Campbell, D. P. .. 483
Campbell, D. T. .. 161
CARMICHAEL, L. .464, 465
CAUL, W. F. 186
CERASO, J. 221
Chall, J. 661
CHAPMAN, N. 438
CLINE, M. G. 41
Cofer, C. N. 319
COFFMAN, T. L. .. 631
COLTHEART, M. ... 250
COOPER, L. M. 124
Crothers, E. J. 482
CROVITZ, H. F. 461
CROWDER, W. F. .. 442

DALLENBACH, K. M.
........318, 469, 491
DAVID, K. H., ... 426
DAWSON, W. W. .. 287
DAY, R. H. 637
Deese, J. 646
DES LAURIERS, G. .. 140
DIAMOND, L. 594
DICKSON, D. 263
DIELMAN, T. E. .. 426
DOCKRILL, F. J. .. 28
Dohrenwend, B. S. . 484
DONNELLY, M. B. . 602
DUKES, W. F. 309

EAGLE, M. N. 421
EPSTEIN, W. 205
Epstein, W. 474
EYMAN, R. K. 520
Eysenck, H. J. 470

Ferguson, G. A. .. 656
FISHER, G. H. ....
........453, 541, 626
Fiske, D. W. 661
FITZGERALD, R. E. . 370
FLAX, S. W. 446
Foster, D. 154
Foulkes, D. 154
FREEMAN, R. B., JR. 302

Gaito, J. 656
Galanter, E. 322
GALLOWAY, A. 430
Gibson, J. J. 150
GILBERT, A. J. 586
GLICK, J. A. 608
GLICKMAN, M. 257
Goldstein, A. C. .. 659
GONZALEZ, R. C. .. 269
Gregory, R. L. 159
GRIFFITHS, E. 438
GRISSOM, R. J. 120
Guion, R. M. 326
GUMENIK, W. E. .. 115
Gump, P. 161
GUNN, W. J. 236
Gustad, J. W. 484

HAHN, W. W. 446
Hake, H. W. 472
Hall, C. S. 156
HALL, J. F. 291
Hall, J. F. 650
Halpern, B. F. 653
HAYES, D. P. 619
HELSON, H. 558
HEPBURN, K. L. .. 398
Hering, E. 163
HILGARD, E. R. ..124, 314
Himwich, H. E. .. 659
HIRSH, I. J. 644
HOFMAN, J. E. 345
HOLMES, D. S. 614
HOLMES, N. K. 269
Howard, I. P. 476
Hunt, F. R. 651
Hyman, R. 658

IRWIN, R. J. 384
ISAACSON, R. L. ... 73
Isaacson, R. L. 478
IVERSON, M. A. ... 468

JARRARD, L. E. ... 66
Jenness, A. 324
JERISON, H. J. 448
JONES, Q. R., JR. .. 287
JUDISCH, J. M. ..105 294
JULIUS, M. 492

Kahn, R. L. 480
KARLINS, M. 631
Katz, D. 480
KATZ, M. S. 405
KENNEDY, A. 430
Ketty, S. S. 659
KINNEY, J. A. S. .. 282
Kirsch, J. 485
Klein, D. 484
KOH, S. D. 171
KOPLIN, J. H. 81
KRAUSKOPF, J. 634

LAMM, H. 631
LANDAUER, A. A. .. 250
LEIBOWITZ, H. W. .
.........105, 263, 294
LEVY, B. A. 28
LEWIS, T. C. 66
LOEB, M. 236
LOIZOS, C. M. 51
Louch, A. R. 655

Luce, R. D. 322
LUDVIGSON, H. W. . 186
LURIA, S. M. 14, 282

MCALLISTER, D. E. . 377
MCALLISTER, W. R. 377
McGee, V. E. 657
McGinnis, R. 322
MCNULTY, J. A. .. 28
MacLeod, R. B. .. 163
MAIER, N. R. F. .. 492
Marin, J. 651
MARKS, L. E. 196
MARSHALL, A. J. .. 370
MELTZER, L. 619
MICHON, J. A. 137
MITCHELL, D. 140
MOATES, D. R. 81
MORGAN, A. H. ... 124
Murray, E. J. 154

NATSOULAS, T. 313
Nelson, A. G. 483
NEWMAN, E. B. .. 144
NEWMAN, S. E. ... 100
Nowlin, V. 146

ODOM, P. B. 391

PAPPAS, B. A. 398
Phillips, L. D. 660
Pick, H. L. 476
PICKENS, R. W. .. 442
PLATT, J. R. 110
POLLARD, S. W. .. 263
PORSOLT, R. D. 384
POSTMAN, L. 465

RABBITT, P. M. 1
Rachman, S. 470
RAND, G. 608
Rau, L. 481
RAY, A. J., JR. 642
Richardson, S. A. .. 484
Ricks, D. F. 146
RIMOLDI, H. J. A. .. 602
ROBY, T. B. 276
Rosenzweig, M. R. . 656
Russell, W. A. 650
Ryan, T. A. 319

SARRIS, V. 304, 331
SCHIFFMAN, H. R. . 229
SCHNALL, M. 213
SCHRODER, H. M. .. 631
Schultz, D. P. 661
Schwartz, R. D. 161
Sears, R. R. 481
Sechrest, L. 161
SEGAL, S. J. 257
SERSEN, F. A. 458
SHAW, W. A. 59
SHICKMAN, G. M. . 128
SIEGEL, M. H. 411
Siegel, S. 660
SILVERSTEIN, A. ... 88
Simon, H. A. 650
SINGER, E. 243, 276
Smith, P. C. 326
Smythies, J. R. 659
SPENCE, K. W. 110
SPITZ, H. J. 462
STABLER, J. R. 95
Staniland, A. C. ... 472
Stein, A. 481
Stephenson, W. 154
STICHT, T. G. 625
Stogdill, R. M. 480
STONE, L. A. 142
Stone, P. 651
SUBOSKI, M. D. ... 398
Sundberg, N. D. ... 470

TAYLOR, A. 95
Teagarden, L. J. ... 485
Templeton, W. B. . 476
THOMAS, J. P. 297
Thompson, R. F. .. 653
THURBER, J. A. ... 492
Tinker, M. A. 661

Van de Castle, R. L. 156
Veldman, D. J. 657
VETTER, R. J. 458
VOEVODSKY, J. 124

WADE, N. J. 637
WAPNER, S. 608
WEBB, B. 115
Webb, E. J. 161
WEINBERG, J. 608
WEINSTEIN, S. 458
WEISSMAN, S. 282
Wessman, A. E. ... 146
Wohlwill, J. F. 474
WONG, R. 140
WOODRUFF, B. 558
WORCHEL, C. 434

YOUSSEF, Z. I. 355

SUBJECTS

(References in *italic* figures are to reviews.)

- Ablation, dorsal-hippocampal, rats, 73-80; hippocampal, and maze learning in rat, 66-72.
- Acknowledgment, an, 318.
- Action, explanation and, *655 f.*
- Activity, apparatus for measuring, 438-445.
- Acuity, visual, change with age, 295 f.
- Adaptation-level, anchor effects and, 334-339, 344; anchor size and, 304-309; comparative method of constant stimuli in, 586-593; step intervals and, 339-344; weighted average model and, 331-344.
- Affection, scales of, in judgment of musical excerpts, 177-184.
- After-effects, figural, Ganz's hypothesis on, 462 ff.
- Age, Ponzo illusion and, 105-109; size-constancy and, 267 f., 294 ff.; size-matching and, 263-268; visual acuity, change with, 295 f.
- Ambiguity, in figures, measurement of, 541-557.
- Ambiguous figures, *see* Puzzle pictures.
- Anchor effects, adaptation-level theory and, 334-339, 344; children's reproductions of changing pictures and, 215-220.
- Animal: Goldfish, resistance to extinction in, 269-275; rat and, in resistance to extinction as function of reward, 269-275.
  - Rat, discrimination-reversal over training and, 363-369; dorsal hippocampal ablation in, passive-avoidance tasks and, 73-80; double alternation learning in temporal maze, 594-560; goldfish and, in resistance to extinction as function of reward, 269-275; hippocampal ablation, maze learning and, 66-72; maze learning in, 66-72; passive-avoidance learning in, 73-80; performance in runway, 95-99.
  - Salamander, Tiger, avoidance-learning in, 642 f.
- Announcement, editorial, 491.
- Antilogarithmic grid, 128-133.
- Apparatus: Activity records, 442-445.
  - Auditory, for group experiments, 120-123.
  - Interpersonal behavior, 42 f.
  - Microphone, bone-conducting, 619-624.
  - Movement monitor, electromagnetic, 438-441.
  - Optical, for calibration of frontal plane, 302 f.; light intensity variator, 297-301.
  - Shock 'scrambler,' solid state, 446 f.
  - Size-distance experiments, 251 f.
  - Stabilimeter, electromagnetic, 438-441.
  - Switch, electronic, for reaction-time measures, 625.
  - Visual, homogeneous field, 405 f.
- Apprehension, span of, fixation-point and, 370-376, in figural aspects, 449-452.
- Arousal, imagery and, 258-262.
- Association, integration and, in serial learning, 355-362; in language and thought, *646-649.*
- Association-hierarchies, individual, predicted from cultural frequencies, 86-94.
- Attensity, interference in color-name test (Stroop) and, 416-420.
- Audition, apparatus for group experiments in, 120-123; loudness, ratio estimation of, 59-65.
- Auditory theory, Meyer's 645.
- Autokinesis, reduction of, with eye-position feedback, 434-437, with restricted viewing, 434-437; set specificity and, 426-429; suggestion and, 426-429; temporal arrangement and, 426-429.

- Behavior, neural mechanism of, *478 ff.*
- Belongingness, verbal recall and, 309-312.
- Bezold-Brücke phenomenon, color naming and, 19 f., 23, 26.
- Biography: Baird, J. W., 491.
  - Bentley, Madison, 491.
  - Bitterman, M. E., 491.
  - Boring, E. G., 491.
  - Dallenbach, K. M., 491.

Hall, G. S., 491.
Hull, Clark, 315 ff.
Newman, E. B., 491.
Postman, Leo, 491.
Sanford, E. C., 491.
Stumpf, Carl, 644.
Titchener, E. B., 491.
Washburn, M. F., 491.
Yerkes, R. M., 315.
See also: Necrology.
Body-tilt, centrifugation and, in changing direction of body-force, 637-639.

Child-rearing, identification and, *481 f.*
Choice-behavior, *660 f.*
Closure, in broken circle, 614-618.
College entrance examination scores, decision-making and, 245-249.
Color-discrmination, stimulus repetition and, 411-415.
Color-naming, Bezold-Brücke phenomenon, 19 f., 23, 26; stimulus-duration and, 14-27; stimulus-intensity and, 14-27.
Color-word test, attensity and interference in, 416-420; pronounceability and interference in, 416-420
Concept-clusters, analysis of, 345-354.
Concept-formation, decision-making and, 245-249.
Concepts, educational, differentiation of, by semantic differential, 348-351.
Conceptual complexity, identification and, 631-634; recall and, 631-634.
Conditioning, extinction in, 317.
Counseling psychology, *483 f.*
Creativity, association in, 493-519; individual differences in, 492, 502-519; information-storage and utilization in, 492-519; measures of, 492-519; mechanisms of, postulated, 492 f., 518 f.; mental competence scores and, 510-519; as reorganization of experience, 492 ff., 518 f.; story writing, as measure of, 495-502, 513, sex differences and, 509 ff.; word image as measure of, 509-513.
Cues, of perspective, effect on size- and distance-judgments, 250-256; visual, binocular vs. monocular in size-matching, 263-268.

Deaf, rule-learning in, 391-397.
Decision-making: Ability and, 243-249.
  College entrance examination (and, 245-248.
  Concept-formation and, 245-249.
  Factor analysis of variables in, 243-249.
  Memory-span and, 245-249.
  Rote learning ability and, 244, 246-249.
  Sex-differences in, 245-249.
  Strategy in, 243-249.
Demonstration, classroom, heterochromatic stabilized images, 634-637.
Depression, feeling of, *146 f.*
Detection, cross-modal, correlation in, 236-242; intensity changes, 236-242.
Deuteranopia, color naming in, duration and, 20 ff.
Discrimination, color, quantal hypothesis and, 412-415, stimulus-repetition and, 411-415.
Distance, estimates of, in air and water, 282-286, underwater, 282-286.
Dreams, content analysis of, *156-159.*
Drive, adversive, instrumental learning and, 377-383; incentive and, in straight alley paradigms, 95-99; reward and, in instrumental learning, 377-383.

Elation, feeling of, *146 f.*
Errata, 45, 318, 469.
Errors, global, in children's reproductions of pictures, 213-220.
Eye-movements, pattern reversals and, 580-583; in perception of moving figures, 575-585.

Feedback, positioning of targets and, 405-410.
Feeling, personal, scales of inadequacy, *149.*
Figures, ambiguous, *see* Puzzle pictures; moving, distortions in, 572-585; reversible, measure of ambiguity in, 541-557, *see* Puzzle pictures.

Game, *JAM,* as isomorph of tic-tac-toe, 137-140, relation to number scrabble, 139 f.
Ganz's hypothesis, figural after-effects, 461 ff.
Generalization, spatial, instructions and, 189-195, in multiple-probability learning, 186-195, individual strategies and, 193 ff.
Grammar, of English sentences and semi-sentences, 196-204.

Happiness, personality characteristics associated with, *148;* variability in, among students, *146 f.*
Hearing, *see* Audition.
Hypnoid syndrome, Perky effect as, 262.

Identification, conceptual complexity and, 631-634.

- Illusion: Apparent movement of a vista, 448-453.
  - Motion, 142 f.; nausea from, 143.
  - Müller-Lyer and Ponzo, common principle relating, 626-631; and Necker cube, compared, 453-457.
  - Necker cube, distortion gradients in, 453-457; Ponzo and Müller-Lyer, compared, 453-457.
  - Perspective, 448-453.
  - Ponzo, age and, 105-109, past theories criticized, 108 f.; Müller-Lyer, common principle relating, 626-631; changes in estimates of, with age, 294 ff.; as a distortion-gradient, 626-631; Müller-Lyer and Necker cube compared, 453-457.
  - Ponzo illusion, age and, 105-109.
  - Size, in air vs. water, 282-286; of familiar objects under differing viewing conditions, 229-235; false perspective cues and, 250-256; of holes, 51-58; underwater, 282-286.
  - Size-distance, 205-212.
  - Target-positions, in homogeneous visual field, 405-410.
- Imagery, arousal level and, 258-262; body position and, 257-262; tension and, 261 f.
- Incentive, drive and, 95-99.
- Industrial psychology, *480 f.*
- Information, storage and utilization of, individual differences in, 492-519.
- Information-processing, conceptual complexity and, 631-634.
- Inquiry, psychological, *658 f.*
- Integration, association and, in serial learning, 355-362.
- Intelligence, decision-making and, 243-249.
- Intervening variables, Spence on, 317.
- Interviewing, *484 f.*
- Invariance hypothesis, an examination of, 250-256; size-distance, 640 f.

- Judgment: Accuracy of, by mental retardates and normals, 536 f.
  - Affective, of musical excerpts, 177-184.
  - Comparative, of musical excerpts, time errors in, 171-185; of weights, time error and, 115-119.
  - Cross-modal, of holes, 51-58.
  - Discriminative, length, by mental retardates vs. naïve and trained normals, 521-529, 532.
  - Distance, 205-212; in air vs. water, 282-286; false perspective cues and, 250-256; underwater, 282-286.
  - Grammaticalness, as a function of inversion of word-order, 197-204; syntax and, 197-204; sentences and semi-sentences, 196-204.
- Looking behavior in another person, 41-50; head movements and, 47-50.
- Loudness, lambda scale of, 59-65; ratio-estimates of, 59-65; sone scale of, 59-65.

- Language, associational structure in, *646-649.*
- Learning, *650 f.:* Ability, rote, decision-making and, 244, 246-249.
  - Alternation, double, in temporal maze, by rat, 594-601; as symbolic process, 595, 599 ff.; training procedures and 594-601.
  - Associations, convergent, in mediated generalization, 81-87; mediation-hypothesis and, 81-87.
  - Chemical approach in, *656 f.*
  - Concept, *651-653.*
  - Discrete-trials lever-pressing paradigm, 110-114.
  - Discrimination, Spence on, 316.
  - Discrimination-reversal, overtraining and, 363-369.
  - Drive and incentive in straight alley paradigms, 95-99.
  - Feigenbaum's theory of, 430-433.
  - Goal response, anticipatory, 95-99.
  - Instrumental, avoidance-conditioning in Tiger salamander, 642 f.; drive, adversive, and, 377-383, and reward in, 377-383.
  - Intertrial interval and response-competition in discrete-trials lever-pressing, 110-114.
  - Maze, hippocampal ablation and, 66-72; intertrial interval and, 66-72.
  - Overtraining-reversal effect (ORE), 363-369.
  - Paired associates, cross-modal stimuli in, 100-104; serial learning and, 356-362; serial position and retrieval time in, 430-433; serial presentation and, 100-104.
  - Passive-avoidance, in dorsal-hippocampectomized rats, 73-80; shock and, in rat, 73-80.
  - Perceptual, *474 ff.;* exposure to stimulus-invariant and, 205-212.
  - Probability, multiple instructions and, 189-195, spatial generalization in, 186-195, as a function of individual strategies, 193 ff.; strategies and, 276-281, *660 f.*
  - Recall, belongingness of elements and,

309-312; confidence ratings and, 400 f.; free, associative organization and, 421-425, instructions and, 429 ff., learning strategies on, 421-425, rehearsal and, 421-425; long and short lists, 221-228; Receiver Operating Characteristic (ROC) as measure of, 400-404; recognition and, compared, 398-404.
  Recognition, confidence ratings and, 400 f.; recall and, compared, 398-404; Receiver Operating Characteristic (ROC) as measure of, 400-404.
  Resistance to irrelevant information in, 1-13.
  Response-competition, intertrial interval and, in discrete-trials lever-pressing, 110-114.
  Reward, resistance to extinction in goldfish and, 269-275.
  Rules of order in the deaf, 391-397.
  Serial, association and integration in, 355-362; in children, facilitated by mediation, 140 ff.; integration and association in, 355-362; linguistic processes in, 133-137; non-linguistic processes in, 133-137; paired associates, and, 356-362; positional cues in, 355; retroactive interference in, 133-137.
  Serial position error (SPE) and, 133-137.
  See also: Association.
Learning theory, mathematical, *482 f.*
Linguistics, judgment of grammaticalness in sentences and semi-sentences, 196-204.
Loudness, scales of, compared, 59-65.

Mastectomy, preferences for body parts after, 458-461.
Mathematical psychology, *322 ff.*
Meaning, perception of physical stimulus structure and, 29-40; subthreshold perception of, 28-40.
Meaningfulness, determinants of, in discrete verbal units, 291-293.
Measurement, non-reactive, *161 f.;* principles of, in the social sciences, *161 f.*
Mediation, role in facilitation of serial learning in children, 140 ff.
Meetings: Amer. Philos. Soc., 464 f.
  Amer. Psychol. Assn., 144 f.
  Eastern Psychol. Assn., 468 f.
  Nat. Acad. Sciences, 465.
  Soc. Exper. Psychologists, 465 f.
  Southern Soc. Philos. and Psychol., 466 f.
  Southwestern Psychol. Assn., 643 f.
Memory, role of hippocampus in, 66-72.
Memory-span and decision-making, 245-249.
Method: Association, free, in language and thought, *646-649.*
  Color-naming, 14-27.
  Cross-modal, in detection, 236-242.
  Interpolated stimuli, relevance on time-error in lifted weights, 115-119.
  Learning, discrete-trials lever-pressing paradigm, 110-114; double alternation, 594-601; paired associates, 100-104, 356-362, 430-433; recall, 221-228, 400 f., 421-425, 631-634, compared with recognition, 400-404; recognition, compared to recall, 400 f.; serial, 133-137, 140 ff., 355-362; signal, detection procedure, 400-404; straight alley paradigm, 195 f.
  Psychophysical, *see* Psychophysics.
  Semantic differential, 345-354.
Meyer, Max F., 644 f.
Microphone, bone-conduction, 619-624.
Mood, personality and, *146-150.*
Motivation, *319-322;* mathematical formulations of, 316 f.; reward, resistance to extinction in goldfish, and, 269-275; *see* also Drive.
Movement-monitor, 438-441.

Necker cube, distortion gradients and, 453-457.
Necrology: Meyer, Max F., 644 f.
  Spence, K. W., 314-318; portrait, 170; acknowledgement, 318.
Neurosis, causes of, *470 ff.*

Organization, visual, in children, 213-220.

Pain, derived from cold, 124-128, formula for, 125, 127 f.; suprathreshold, measurement of, 124-128.
Perky effect, body position and, 257-262.
Personality, factor analysis of, flaws in, *150;* mood and, *146-150.*
Perception: Adaptation-level, anchor-size and, note on, 304-309.
  Anorchoscopic, 572-585.
  Auditory, loudness, 384-390.
  Binaural tone, loudness as function of bandwidth, 384-390.
  Closure, in gapped circle, 614-618.
  Color, interference in Stroop test, 416-420; guantal hypothesis and, 412-415; stimulus-repetition and, 411-415.

  - Cross-modal, of holes, 51-58.
  - Distance, estimates in air vs. water, 282-286; false perspective cues and, 250-256; retinal size and, 205-212; underwater, 282-286.
  - Figural after-effects, Ganz's hypothesis and, 462 ff.
  - Figure, ground and, phenomena, 549, 554 f.; of moving forms, distortions in, 572-585, eye-movement in, 575-585; tilt-distortion in, 578 ff.; multi-aspect, 550 ff.; reversible, chalice and faces, 546, measure of ambiguity in, 541-557.
  - Fluctuations of, in reversible figures, 547.
  - Ganz's hypothesis on figural after-effects, 461 ff.
  - Imagery, body position and, 257-262; tension and, 261, f.
  - Intensity-changes, 236-242.
  - Invariance hypothesis, 250-256.
  - Letters, dependent on relation to fixation-point, 370-376.
  - Looking behavior in another person, 41-50, head movements and, 47-50.
  - Loudness, binaural summation in, 384-390, lambda scale of, 59-65, sone scale of, 59-65.
  - Pattern reversals, eye-movements and, 580-583.
  - Perky effect, body position and, 257-262; hypnoid syndrome and, 267.
  - Sense, *150-154. See* also: Sensation.
  - Size, age, distance and, 263-268; binocular vs. monocular cues in, 263-268; estimates, in air vs. water, 282-286; false perspective cues and, 250-256; familiar objects, under different viewing conditions, 229-235; of holes, 51-56; monocular vs. binocular cues in, 263-268; underwater, 282-286.
  - Size-constancy, age and, 294 ff.; distance and, 294 ff.
  - Size-distance, hypothesis, underwater, 282-286.
  - Spatial orientation, *476 ff.*
  - Subliminal, of meaning, 28-40.
  - Tactual, of hole size, 52-58.
  - Tactual-kinesthetic, speed of movement on, 608-613.
  - Visual, *159 ff.;* in children, 213-220; Hering on, *163;* mechanisms of, *159 ff.;* of hole size, 52-58; *see* also Vision.
- Perceptual ecology, *150 f.*
- Perspective, cues of, effect on size and distance judgments, 250-256.
- Physiological psychology, *653 f.*
- Point of subjective equality, shift of, effect of interpolated weights on, 115-119.
- Preference, 'background effect' and, 177-181; body part, after mastectomy, 458-461; comparisons of, for musical excerpts, time-error in, 171-185; ratings of, for musical excerpts, 171-185.
- Probability learning, strategies in, *660 f.*
- Problems, creative, logical solutions and, 461 f.
- Programming, computer, *657 f.*
- Psychology of the 'other one,' Meyer's formulation of, 644 f.
- Psychophysics: Method, constant stimuli, comparative adaptation-levels (CAL), 586-593, judgments in, 586-593; constant sum, 533-540; interpolated stimuli, 115-119; paired comparison, 521, 525 ff, 529-532.
  - Point of subjective equality (*PSE*), interpolated weights and, 115-119.
  - Quantal hypothesis in color perception, 411-415.
  - Scales, category, 602-607; constant sum, 527 ff., 532-540; discriminative, effect of sophistication on, 520-540; loudness, 59-65; magnitude, 602-607; paired comparison, 521, 525 ff., 529-532, and constant sum, compared, 533-540; ratio, effect of sophistication on, 520-540.
  - Scaling, by mental retardates, compared with trained and naive normals, 520-540.
  - Time-error in, 171 f.
- Puzzle pictures, gypsy and girl, 543; hunter and rabbit, 545; husband and father-in-law, 545; lion and rose, 545; measures of ambiguity in, 541-557; slave market and Voltaire, 555; snail and elephant, 545; wife and mother-in-law, original and negative, 547.

- Ratings, of preferences for musical excerpts, 171-185.
- Reaction-time, apparatus for study of, 251.
- Reading, *661 f.*
- Recall, conceptual complexity and, 631-634.
- Redundancy, patterns of, *472 ff.*
- Retardates, mental, ratio scaling by, compared with trained and naïve normals, 520-540.
- Reversible figures, ambiguity in, 541-557; *see* also: Puzzle pictures.
- Reward, drive and, in instrumental learning, 377-383.

Scale, affective, in judgment of musical excerpts, 177-184; loudness, lambda, 59-65, ratio-estimates of, 59-65, sone, 59-65; pain, suprathreshold, 124-128; semantic differential, 345-354.
Scaling, *see* Psychophysics, scaling.
Schizophrenia, amines and, *659 f.*
Self, locus of, determined by finger tracings on head, explanation, 313 f.
Semantic differential, analysis, of concept-clusters by, 345-354, of educational concepts by, 348-351.
Sensation, auditory, detection of correlated with visual detection, 236-242; cold, painful sensation from, 124-128; pain, measurement of, 124-128; tactile, clues for magnitude of, 287-290; visual, detection of, correlated with auditory detection, 236-242; *see* also: Perception.
Sense-perception, a functional psychology of, *150-154.*
Senses, as perceptual system, *150-154.*
Sensory deprivation, *661.*
Sex, decision-making as a function of, 245-249.
Shakespeare, psychoanalytic study of, *485 f.*
Shock, apparatus to deliver, 446 f.
Size, estimates, in air vs. water, 282-286, underwater, 282-286.
Size-constancy, age and, 267 f., 294 ff.; distance and, 294 ff.; underwater, 284-286.
Size-distance hypothesis, underwater, 282-286; invariance hypothesis, 640 ff.; judgment, apparatus for study of, 251 f.
Size-matching, age and, 263-268; binocular vs. monocular cues in, 263-268; distance as function of, 263-268; monocular vs. binocular cues in, 263-268.
Sleep, arousal and, *154 ff.;* dreams and, *154 ff.;* as a drive, *154 f.;* eye-movements in, *155 f.;* psychology of, *154 ff.*
Social psychology, *324 ff.*
Solutions, logical, 461 f.
Spatial orientation, *476 ff.*
Spence, K. W., 314-318; portrait, facing 171.
Stabilized images, heterochromatic, classroom demonstration of, 634-637.
Story-writing, as measure of creativity, 495-502, 513.
Strategy, games, tic-tac-toe vs. *JAM,* 137-140; probability learning and, 276-281, *660 f.*
Stroop test, alternation and interference in, 416-420; pronounceability and interference in, 416-420.
Summation, binaural, of tones as function of bandwidth, 384-390.
Syntax, in judgment of grammaticalness of sentences, 197-204.

Testing, of personnel, *326.*
Therapy, behavioral, *470 ff.*
Thought, associational structure in, *646-649.*
Time-error, in comparative judgment of weights, interpolated stimuli and, 115-119.
Torque, sensitivity to, knob radius and load in, 558-571, Weber's law and, 567-571.

Unhappiness, personality characteristics associated with, *148.*

Vision, color-naming, 19-27; defects, deuteronopia and color-naming, 20 ff.
Visual feedback, accuracy of target positioning and, 405-410.
Visual figures, changing, children's reproductions of, 213-220.
Visual search, irrelevant symbols and, 1-13; speed of, 1-13.

Word-order, inversions of, and judgment of grammaticalness, 197-204.
Word-usage, as measure of creativity, 509-513; mental competence scores and, 510-519.

BUREAU OF EDNL. & PSYL. RESEARCH. LIBRARY